Journal of Applied Psychology

Edited by Donald G. Paterson, University of Minnesota

Consulting Editors

George K. Bennett, *Psychological Corporation*
Walter V. Bingham, *Washington, D. C.*
Harold E. Burtt, *Ohio State University*
Allen L. Edwards, *University of Washington*
Clifford E. Jurgensen, *Minneapolis Gas Co.*
Irving Lorge, *T. C. Columbia University*
Quinn McNemar, *Stanford University*
Alexander Mintz, *City College of New York*
James P. Porter, *Danville, Illinois*
Julian B. Rotter, *Ohio State University*
Edward K. Strong, Jr., *Stanford University*
Donald E. Super, *T. C. Columbia University*
Morris S. Viteles, *University of Pennsylvania*
Alfred C. Welch, *Knox-Reeves, Minneapolis*

Table of Contents

Is How Supervise? an Intelligence Test?: K. A. Millard 221

Reading Ease Scores for File's How Supervise?: P. W. Maloney 225

Temperament Traits of Executives and Supervisors Measured by the Guilford Personality Inventories: J. S. Guilford 228

Psychological Test Performance of Steel Industry Production Supervisors: W. A. Poe and I. A. Berg 234

A Cluster Analysis of Office Operations: L. L. Thomas 238

A Factor Analysis of a Salary Job Evaluation Plan: A. H. Howard and H. G. Schutz 243

A Factor Analysis of Employee Attitudes: R. J. Schreiber, R. G. Smith, Jr., and T. W. Harrell 247

A Factor Analysis of Terman and Miles' M-F Test: C. F. Ford, Jr. and L. E. Tyler 251

Vocational Interests of Retired YMCA Secretaries: W. A. Verburg 254

Predicting Success in Law School: R. F. Berdie and W. L. Layton 257

The ACE Psychological Examination and High School Standing as Predictors of College Success: N. Frederiksen and W. B. Schrader 261

Education and Prediction of Military School Success: V. Zachert and A. S. Levine 266

The Use of Levers in Making Settings on a Linear Scale: W. L. Jenkins and M. W. Olson 269

Dimensional Analysis of Motion: III. Complexity of Movement Pattern: G. Rubin, P. Von Trebra, and K. U. Smith 272

Book Reviews 277

New Books, Monographs, and Pamphlets 284

American Psychological Association

Vol. 36, No. 4 **August, 1952**

Journal of Applied Psychology

Published Bi-monthly by the American Psychological Association, Inc.
Prince and Lemon Sts., Lancaster, Pa.

Annual subscription, $6.00; single copies, $1.25

Subscriptions and business communications should be sent to

American Psychological Association
1515 Massachusetts Avenue N.W.
Washington 5, D. C.

Articles for publication and books for review should be sent to the Editor

Professor Donald G. Paterson, Department of Psychology
University of Minnesota, Minneapolis 14, Minnesota

This Journal gives prompt consideration to manuscripts reporting *original investigations* in any field of applied psychology except clinical and consulting psychology. A descriptive or theoretical article is occasionally accepted if it deals in a distinctive manner with a problem of applied psychology. The policy is, however, to favor papers dealing with quantitative investigations of direct value to psychologists working in the following fields: Vocational diagnosis and occupational guidance; educational diagnosis, prediction and guidance at the secondary school level and higher; personnel selection, training, placement, transfer and promotion in business, industry and government service including the armed forces; supervisory training in business, industry and government; bio-mechanics or design of machines to fit the human operator; illumination, ventilation and fatigue in industry; job analysis, description, classification and evaluation; measurement of morale of executives, supervisors, or employees; surveys of opinion on social or political issues, such as those conducted by The Psychological Corporation; psychological problems in market research and in advertising.

Articles may be under 500 words. The maximum is 12,000 words, the average in the neighborhood of 4,000 words. To reduce lag of publication, adherence to the rule of "brevity consistent with clarity" is encouraged.

A lapse of six to twelve months occurs between acceptance of an article and its publication, the lag varying with the rate at which manuscripts are submitted. If, however, an author is prepared to defray the costs of printing the necessary extra pages, he may arrange for earlier publication without thereby postponing the appearance of manuscripts by other contributors. This enables the management to provide space in addition to the scheduled 64 pages per issue. "Early publication" is thus a direct contribution to the subscribers. By cutting down lag in publication, it also benefits those authors whose articles are published in regular turn.

Tables, footnotes and references as well as text of manuscripts should be typed double-spaced throughout. Authors should adhere to the conventions described by J. E. Anderson and W. L. Valentine in "The preparation of articles for publication in the journals of the American Psychological Association," *Psychol. Bull.*, 1944, 41, 345–376. A reprint of this article will be loaned to any prospective contributor who does not find it in his library.

Entered as second-class matter, August 19, 1943, at the post office at Lancaster, Pa., under the act of March 3, 1879

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized October 10, 1947

Copyright, 1952, by The American Psychological Association, Inc.

Journal of Applied Psychology

VOL. 36, NO. 4 AUGUST, 1952

Is *How Supervise?* an Intelligence Test? *

Kenneth A. Millard

Macalester College

The conviction that *How Supervise?* "is in large part merely a test of intelligence" was expressed by Slocombe (**10**) before any evidence on the point was available. To this date apparently only one published report of the relationship between *How Supervise?* and a mental ability test has appeared. And that report was badly misinterpreted in a later article that commented upon it. At the same time, Harrell (**4**, p. 137) and especially Mosier (**6**) have echoed Slocombe's conviction.

In an aircraft plant Sartain (**9**) obtained a product-moment correlation of $-.44$ (significant at the 1% level) between the *Adaptability Test*, Form A and *How Supervise?*, Experimental Edition, Form A for a group of 40 supervisors. File (**1**) had reported previously a correlation of .35 between highest educational level reached and *How Supervise?* for a group of 577 factory supervisors. Following publication of Sartain's findings, File and Remmers (**2**) attempted to explain what they considered the marked discrepancy of these two studies. But apparently File and Remmers overlooked a parenthetical statement in Sartain's article that removes any necessity for explaining a discrepancy since none exists.

Sartain, in commenting upon his findings, wrote: "The correlation between *Adaptability* and *How Supervise?* indicates that general mental ability goes with favorable supervisory attitudes (low scores on this test indicating a favorable attitude) to a moderate degree" (**9**, p. 331). The published method of scoring the test is such that high scores, rather than low scores (as indicated by Sartain), represent "favorable" supervisory attitudes. For some unknown reason Sartain apparently used a "reverse" method of scoring, and actually his findings and those of File are in close agreement.

Procedure

The present study obtained further evidence on the relation between *How Supervise?* and measures of intelligence. Three different kinds of supervisors were included. Of the 49 factory supervisors, all men, 28 worked in a flour mill and 21 in a textile manufacturing plant. Eighteen of the 71 office supervisors were employed in the home office of a life insurance company. The other 53 were distributed among the accounting, advertising, circulation, news and features, and promotion departments of a metropolitan newspaper. Twelve of the office supervisors were women. These factory and office supervisors were all from the lower levels of management. Very few of them supervised only other supervisors, although many of them had responsibility over some other supervisors as well as directly over non-supervisory employees.

The third group, also employed by the newspaper, supervised carriers and dealers. This group of 77 included 51 city district managers, 18 country district managers, and eight zone supervisors (who supervise only district managers). Seventeen of the city district managers were women; all others in this total group were men.

Each of the supervisors in all three groups was administered *How Supervise?*, Form A and the *Adaptability Test*, Form B. Both of these

* This article is based on the findings reported in the writer's thesis submitted in partial fulfillment of the requirements for the Ph.D. degree in psychology at the University of Minnesota. The data reported here on supervisors of newspaper carriers and dealers were not included in the thesis, although gathered at the same time. The writer wishes to thank his major adviser, Professor D. G. Paterson, for important guidance throughout the planning and conduct of the study.

were answered anonymously, since an opinion questionnaire about company policies and practices was one of the other measures used. So that all measures for the same individual could be compared, a code number was used for each individual.

For each of the three groups of supervisors, the product-moment correlation between *Adaptability Test* and *How Supervise?* scores was obtained. (No significant sex differences in either mean or variance for either test were found, so data for men and women were combined.) Calculations were made from grouped data. The assumption of rectilinearity of regression was tested for each supervisory group. In no case was the deviation from rectilinearity sufficient to warrant rejection of the assumption. The only test of the assumption of homoscedasticity was the subjective one of visual inspection of the scattergram.

Because of the known substantial relationship between tested intelligence and highest educational level reached, this study also determined the relationship between the latter and *How Supervise?* scores. Each supervisor, on a personal information questionnaire, checked one of the five following categories: 8th grade or less, attended high school but didn't graduate, graduated from high school, attended college but didn't graduate, graduated from college. Then, for each of the three groups of supervisors, analysis of variance was applied to determine whether there was significant difference in mean *How Supervise?* scores among the educational level groups. (Using the method of Welch, the assumption of homogeneity of within-groups variances was tested and in each case found to be a tenable assumption.) Where analysis of variance revealed significant difference among these means, epsilon, the unbiased correlation ratio (**8**), was computed to get an indication of the degree of relationship between highest educational level reached and *How Supervise?* scores. If the assumption of rectilinearity of regression is warranted, then epsilon is comparable to the product-moment correlation coefficient.

Results

Factory Supervisors. The relation between *Adaptability Test* and *How Supervise?* scores is presented in Table 1. The r of .71 for the factory supervisors is not much less than the equivalent form (Forms A and B) reliability of .77 for *How Supervise?* (**3**). The 99% confidence limits for this r of .71 are .47 and .85. For our particular group of factory supervisors, 50 per cent of the variance in *How Supervise?* scores is accounted for by the variance in *Adaptability Test* scores.

Analysis of variance revealed a significant (1% level) relationship between highest educational level reached and *How Supervise?* for the factory supervisors. The value of epsilon, computed to determine the degree of this relationship, was .45. This value is similar to the correlation of .35 reported by File (**1**) for the same relationship.

Office Supervisors. The r between *Adaptability Test* and *How Supervise?* scores for the office supervisors is .22, which is significant only at the 8% level. Moreover, the relationship between highest educational level reached and *How Supervise?* for this group of supervisors is not significant at even the 5% level.

Carrier and Dealer Supervisors. For the supervisors of newspaper carriers and dealers the r between *Adaptability Test* and *How*

Table 1

Relation Between *Adaptability Test* and *How Supervise?*

Supervisory Group	N	*Adaptability* Mn	*Adaptability* SD	*How Supervise?* Mn	*How Supervise?* SD	r: *Adaptability* and *How Supervise?*
Factory	49	12.9	6.7	42.0	12.0	.71**
Office	71	20.9	6.0	45.7	8.8	.22
Newspaper carriers and dealers	77	16.9	6.9	42.1	10.3	.62**

** Significant at the 1 per cent level.

Supervise? scores is .62, a value significant at the 1% level. The relation between highest educational level reached and *How Supervise?* for this group is significant at the 5% level, with the degree of the relationship being indicated by an epsilon of .30.

Discussion

It is pertinent to compare our r of .71 for factory supervisors with Sartain's (**9**) r of .44 (giving it a plus sign to have a comparable meaning) between the same two measures for a group of aircraft plant supervisors. The difference between these two r's reaches significance at the 6% level. Whether our larger r might be at least partly a function of greater variability for either or both measures is unknown because Sartain did not report any measures of variability. It might be that the experimental edition of *How Supervise?*, which he used, was not as much a measure of intelligence as the later editions.

The findings from the present study focus particular concern on the substantially lower relationship between intelligence and *How Supervise?* for the office supervisors than for either of the other two groups of supervisors. (The differences between the office supervisors' r and that for each of the other two groups are significant at considerably less than the 1% level; on the other hand, the difference between the r's for the other two groups does not even approximate significance at the 5% level.)

To seek an explanation of this lower r for the office supervisors we first take a look at the variabilities of the three groups on the two measures (cf. Table 1). (Incidentally, in the case of *How Supervise?* the means and variabilities for our groups of supervisors approximate the values reported in the test manual for comparable groups. For the *Adaptability Test* the test manual does not give data for any directly comparable groups.) Comparing the variances on the *Adaptability Test* for each of our pairs of supervisory groups, we find none of the F values significant at even the 10% level. Two of the three pairs do not have significant differences for variance on *How Supervise?* either. However, the difference in variance on *How Supervise?* between the factory and office supervisors is significant at the 2% level. Therefore, we know that at least part of the reason for the smaller r for the office group is its smaller variability. To determine the effect of this smaller variability on the r for the office group, a new r was computed, using the *How Supervise?* SD of the factory group as the uncurtailed SD for that measure. This new r for the office group was .29, a value just barely short of significance at the 5% level.

Since differences in variability account for very little of the difference in r's we next look at differences in means. On the *Adaptability Test* all three differences between means are significant at probabilities of 1% or less. In the case of *How Supervise?* only one of the differences between means is significant; viz., that between the office supervisors and the carrier and dealer supervisors, which is significant at the 5% level. So we find no consistent pattern between differences between means on the one hand and differences between r's on the other. (In determining significance of difference between means the t test was used in all instances but one. For the difference between means of office and factory supervisors on *How Supervise?*, t is not appropriate because of the significant difference in variances. Therefore, in that case the Behrens-Fisher d test (**5, 11**) was used.)

Another relevant comparison is that for highest educational level reached. The distribution for each supervisory group on this variable is presented in Table 2. Each of the pairs of distributions was compared by means of the chi-square test. For each of the three comparisons the chi-square value was significant at the .1% level, a finding consistent with the significance of differences between means on the *Adaptability Test*.

It is seen then that neither differences in variability nor differences in means provide any adequate explanation for the markedly lower r of the office supervisors. There remains one additional characteristic which warrants examination; viz., the relative variability on each measure for the three groups. The use of the coefficient of relative variation (the ratio of SD to the Mn, expressed as a percentage) with psychological measurements has been questioned because of the lack of a true zero point and unequal units of measure-

Table 2

Frequency Distributions for Highest Educational Level Reached

Educational Level	Supervisory Group		
	Factory*	Office	Newspaper Carriers and Dealers*
8th grade or less	18	2	5
Attended high school but didn't graduate	15	8	6
Graduated from high school	7	14	45
Attended college but didn't graduate	4	19	17
Graduated from college	4	28	3

* One factory supervisor and one carrier and dealer supervisor did not give information about highest educational level reached.

ment. However, Peatman (**7**, p. 171) contends that it is legitimate to compare the coefficients of relative variation of two groups on the *same* psychological measure. He also presents a test of significance for such a difference.

When we compare the coefficients of relative variation on the *Adaptability Test* for our three groups, we obtain these interesting findings: the difference between the office group and each of the other two groups is significant at the 1% level, while the difference between the other two groups is not significant at even the 5% level. A similar pattern shows up when we compare the coefficients of relative variation on *How Supervise?*. The office vs. factory difference is significant at the 1% level, the office vs. carriers and dealers difference is significant between the 5% and 6% level, and the factory vs. carriers and dealers difference does not even approximate significance at the 5% level. It is clearly seen, then, that the pattern of differences between coefficients of relative variation for both measures corresponds with the pattern of differences between *r*'s.

Summary

1. Using the *Adaptability Test* as a measure of intelligence, this study found a substantial correlation between intelligence and *How Supervise?* for factory supervisors and supervisors of newspaper carriers and dealers. For office supervisors, this correlation was considerably smaller and of less certain significance.

2. Using highest educational level reached as an indirect measure of intelligence, similar relationships were found, but of a smaller degree.

3. Only a very small part of the markedly lower relationship between intelligence and *How Supervise?* for the office supervisors is accounted for by differences in variability. Nor do differences in means show a pattern similar to the differences in intelligence-*How Supervise?* relationship for the three groups of supervisors. However, the pattern of differences in *relative* variability corresponds precisely with the pattern of differences for this relationship.

4. The frequently expressed conviction that *How Supervise?* is essentially an intelligence test seems to be considerably substantiated for two kinds of supervisors, but not for a third, by this study.

Received August 23, 1951.

References

1. File, Q. W. The measurement of supervisory quality in industry. *J. appl. Psychol.*, 1945, **29**, 323–337.
2. File, Q. W., and Remmers, H. H. Studies in supervisory evaluation. *J. appl. Psychol.*, 1946, **30**, 421–425.
3. File, Q. W., and Remmers, H. H. *Manual for How Supervise?* (Rev. Ed.) New York: The Psychological Corporation, 1948.
4. Harrell, T. W. *Industrial psychology.* New York: Rinehart, 1949.
5. Johnson, P. O. *Statistical methods in research.* New York: Prentice-Hall, 1949.
6. Mosier, C. I. Review of *How Supervise?* In O. K. Buros (Ed.), *The third mental measurements yearbook.* New Brunswick: Rutgers University Press, 1949. Pp. 727–728.
7. Peatman, J. G. *Descriptive and sampling statistics.* New York: Harper, 1947.
8. Peters, C. C., and Van Voorhis, W. R. *Statistical procedures and their mathematical bases.* New York: McGraw-Hill, 1940.
9. Sartain, A. Q. Relation between scores on certain standard tests and supervisory success in an aircraft factory. *J. appl. Psychol.*, 1946, **30**, 328–332.
10. Slocombe, C. S. Appraisal of Mr. File's study. *Person. J.*, 1946, **24**, 251–254.
11. Sukhatme, P. V. On Fisher and Behrens' test of significance for the difference in means of two normal samples. *Sankhya: The Indian J. Statistics*, 1938, **4**, 39–48.

Reading Ease Scores for File's *How Supervise?*

Paul W. Maloney

University of Minnesota

File (**2**), in building *How Supervise?*, defined supervisors as "individuals between the group leader and departmental supervisor levels." As Bellows (**1**) and Tiffin (**7**) have shown, such a group has a wide range of reading ability. Therefore, if these people are to be tested for supervisory ability and nothing else the directions and test items should be highly readable. This point was recently clarified by Johnson and Bond (**5**) in relation to vocational guidance tests commonly given at V. A. Advisement Centers.

File (**2**) recognized this. He said that the items ". . . must be simply worded so that *any* supervisor can see the problem involved" (italics File's). His criterion of readability, however, was merely the opinion of various industrial experts, nine-tenths of whom thought the test to be clearly written. A more objective criterion would be the application of Flesch's (**3**) reading ease formula to the directions and test items. Such an analysis is reported here. The total content (rather than random samples) of the directions and items of Forms A and B has been subjected to the formula.

Results

The results of the Flesch analysis are summarized in Table 1. Flesch's (**4**) grade level equivalents are included. Note that three of the sections of test items are written for individuals with 10th to 12th grade reading ability. Two sections and the directions are at the high school graduate level. One section has a score of 50, lying just between the two above mentioned levels. The mean readability for both forms, directions included, is of high school graduate difficulty. If reading ability is not to be a factor in the response made to the test, it seems safe to conclude that those who are tested with *How Supervise?* should have high school graduate reading ability.

The obvious question, then, is as to whether or not this is the case. Do most industrial supervisors have the reading ability requisite to taking *How Supervise?* A person's reading grade level can be roughly related to his educational level (very roughly—Johnson and Bond (**5**) caution that ". . . an individual's reading level may vary as much as six grades from his educational level"). The educational attainment of industrial foremen is listed in Table 2. This information is from a national sample (5 per cent of the population) drawn in the 1940 U. S. census.

It is apparent that the range of educational attainment of industrial supervisors is wide. The median number of years of schooling

Table 1

Reading Ease of File's *How Supervise?*

Form	Section	Average Sentence Length	Average Number of Syllables per 100 Words	Reading Ease Score	Flesch's Grade Level
A	1	11.3	169.8	52	Some high school
A	2	12.8	184.2	38	High school or some college
A	3	15.6	160.5	56	Some high school
B	1	13.4	165.1	54	Some high school
B	2	12.9	182.9	39	High school or some college
B	3	14.4	168.4	50	High school
Both/Directions		14.3	170.6	48	High school or some college
Means for Total Test		13.5	171.6	48	High school or some college

Table 2

Education Attainment of Male Industrial Supervisors*

Grade Reached	Foremen	
	Number	Per Cent
College:		
4 years or more	12,960	2.64
1 to 3 years	26,640	5.42
High School:		
4 years or more	81,180	16.51
1 to 3 years	102,600	20.87
Elementary School:		
7 and 8 years	197,840	40.24
5 and 6 years	41,220	8.38
Less than 5 years	25,320	5.15
Schooling not reported	3,860	.79
Totals	491,620	100.00

* *Population, The Labor Force, Occupational Characteristics.* U. S. Bureau of Census, 16th Census; pp. 59 and 113.

completed is only 8.8. Less than a quarter of the foremen in the sample had a high school education. We noted that a large part of *How Supervise?* is at the high school graduate reading level. Reading ability, insofar as it is related to educational attainment, is a factor which would be likely to influence the scores of most industrial supervisors when taking the test.

Evidence of this contamination may be found in a study by Millard (**6**). He compared scores on *How Supervise?* with scores on the *Adaptability Test*. For office (higher level) supervisors there was a low correlation, .22. However, the correlation for a group of factory (lower level or first line) foremen was .71. This difference appears to be a function of reading ability. It is unlikely that many of the office supervisors had much difficulty reading *How Supervise?*. But within the group of factory foremen there probably were many who had difficulty understanding material written for people of high school graduate ability. Their scores would then reflect the extent to which they could comprehend the material, rather than their knowledge of the principles of supervision. The *Adaptability Test* also measures the ability to understand written material. So for factory foremen of limited schooling it is only natural that *Adaptability Test* scores should correlate highly with *How Supervise?* scores.

Readability as measured by Flesch is a function of the numbers of syllables and words. *How Supervise?* needs revision in both respects. For example, item 51 in Form A now reads "The only important requirement of a good supervisor is a complete understanding of the jobs he is to supervise." Sentence and word length are unnecessarily long. A shorter version might read "All that supervisors need really know are the jobs they supervise." Similarly, "The goals of management and labor are directly opposed and must always be in conflict with each other" (item 69, Form A) would carry the same connotation if more simply phrased, "Management and labor have different aims which will always conflict." Again, "Constant demands upon the time of top-executives make it impractical for them to spend any time in actual conversations with workers" (item 63, Form B) could just as well be "Top managers are so busy that they should not be asked to talk to the workers."

Summary

The readability of the directions and items in File's *How Supervise?* is at the difficult level. Much of the material is at the high school graduate level of readability. In contrast, less than a quarter of a nation-wide sample of foremen were high school graduates. To the extent to which educational attainment is an indication of reading ability the test can be read only with great difficulty by the average factory foreman. Therefore, for lower level personnel *How Supervise?* is of doubtful validity as a measure of supervisory ability. Three items were rephrased to show how the test could be made more readable.

Received December 5, 1951.
Early publication.

References

1. Bellows, R. M. *Psychology of personnel in business and industry.* New York: Prentice Hall, 1949.
2. File, Q. W. The measurement of supervisory quality in industry. *J. appl. Psychol.*, 1945, **29**, 323–337.
3. Flesch, R. A new readability yardstick. *J. appl. Psychol.*, 1948, **32**, 221–233.

4. Flesch, R. *How to test readability.* New York: Harper and Bros., 1951.
5. Johnson, R. H., and Bond, G. L. Reading ease of commonly used tests. *J. appl. Psychol.*, 1950, **34**, 319–324.
6. Millard, K. A. Is *How Supervise?* an intelligence test? *J. appl. Psychol.*, 1952, **36**, 221–225.
7. Tiffin, J. T., and Colby, A. N. The reading ability of industrial supervisors. *Personnel*, 1950, **27**, 156–159.

Temperament Traits of Executives and Supervisors Measured by the Guilford Personality Inventories *

Joan S. Guilford

University of Southern California

Because of the dearth of effective selection devices for screening of executives and supervisors in industry, this study was conducted as an attempt to uncover traits of temperament which might be important in the successful performance of duties by men in these positions. The meager information concerning the use of temperament tests in connection with executive and supervisor selection makes it difficult to evaluate these devices for selection of such personnel. However, since the pencil-and-paper test is the most available medium at present for the assessment of temperament traits, any investigation of its effectiveness would appear to be worthwhile.

Previous studies have consistently revealed low validity for scores on personality tests in vocational uses. In several cases, however, there has been some improvement in prediction of success and any improvement in prediction, particularly at these high levels in the organizational hierarchy, is valuable It is generally agreed by authorities in the field of personnel selection that traits of temperament contribute greatly to the success or failure of men in executive and supervisory positions, while technical skills which are more easily evaluated are relatively less important. Because of the necessity for well-qualified individuals at these high levels and the relatively large contribution of temperament traits to their success, there is great need for research in the field of personality testing in the industrial situation.

The Problem

The purpose of this study was twofold. It was desired to determine: (1) in what specific respects the executive differs from the supervisor with regard to certain personality traits and (2) the validity of these same traits as related to a criterion of job success, defined in terms of ratings of "job performance."

* This article consists of portions drawn from an unpublished Master's thesis by the writer (1).

The Procedure

The subjects of this study were: (1) 208 executives in a large chain grocery[1]; and (2) 143 supervisors in the same company. The executive in this organization serves as a link between branch headquarters and the stores. He is responsible for supervising the store managers of from twelve to twenty stores in his district, promoting sales, controlling expense consistent with maximum sales, and developing and maintaining competent organizations in his stores. He employs, disciplines, and separates employees in his group or is responsible for recommending such action. The supervisor (or foreman) directs the employees in his plant or department, working with them at all times. He is responsible for the quantity and quality of production in his department, for maintaining satisfactory working conditions, making decisions regarding production problems, handling employee grievances, making suggestions for improvements in the operations he supervises, scheduling and training employees, and employing, disciplining, and separating workers under him or making recommendations for such action.

To further differentiate between these groups, it might be added that in choosing a type of work the district manager (executive) was originally attracted to the "retail culture" which is assumed to emphasize gregariousness and ascendance; the foreman (supervisor) was attracted to the "factory and warehouse culture" involving physical labor and more tangible and obvious duties. The executive is "on his own" in a geographic territory and makes many independent decisions. He provides leadership for his area. The store managers whom he supervises are at a level equivalent to (and in most cases higher than) the foreman group. He typically has 100 to 150 employees in his district. The foreman (or supervisor) usually supervises 20 to 30 men in the unskilled or semi-skilled labor classifications. He operates a department or unit in a factory or

[1] The writer wishes to thank Mr. Glenn E. Mitchell of the Kroger Company for furnishing the data and supplementary information used in this study.

warehouse and is himself under fairly direct supervision.

The tests administered to these two groups were the Guilford series of personality inventories: Inventory of Factors STDCR (**2**); Guilford-Martin Inventory of Factors GAMIN (**3**); and Guilford-Martin Personnel Inventory (**4**). The traits represented in these tests are: (1) Social Introversion-Extraversion; (2) Thinking Introversion-Extraversion; (3) Depression; (4) Emotional Stability; (5) Rhathymia or Impulsiveness; (6) General Activity; (7) Ascendance or Social Boldness; (8) Masculinity of Attitudes and Interests; (9) Inferiority Feelings; (10) Nervousness; (11) Objectivity; (12) Agreeableness; and (13) Cooperativeness.

These factors emerged from several previous factor analyses and so might be assumed to represent relatively distinct dimensions of personality. The reliabilities of the scores have been found to be adequate, ranging generally from .80 to .92.

The criterion was a composite of ratings of "job performance" made, in the case of the executives, by members of the training staff. The raters were two in number and the ratings were made independently. In the case of the supervisors, ratings were made by individuals occupying a position midway between the supervisor and the executive, the superintendents. These ratings were in letter form: an "A" rating was defined as "very outstanding performance," "B" as "good performance," "C" as "average performance," and "D" as "just passing performance." In both groups the men were experienced workers and were rated while on the job. For the sake of computation these letter ratings were converted into numerical ratings of 30, 20, 10, and 0 for A, B, C, and D, respectively. Since the executive's ratings were presented separately, it was possible to compute their intercorrelation which proved to be .53. Assuming this to be a reliability index, the combination of the ratings attains a Spearman-Brown reliability estimate of .69. In the case of the supervisors only the pooled rating was available.

It is difficult to determine the extent to which the testing conditions might have controlled extraneous sources of variance since the tests were administered in branches of the company scattered throughout the East and Middlewest. The greatest danger, of course, lies in the possibility of uncontrolled motivation which might lead to falsification of responses on the test items. In this particular case, the tests were given experimentally so that tendency to falsify is probably much smaller than it would be were these tests used as selection devices. This is one of the most serious weaknesses of these tests, that, with sufficient psychological insight, an individual can alter his profile of traits to fit what he believes should be typical of the individual who successfully fills the position for which he is an applicant.

Results

The first statistical procedure was to determine means and standard deviations for the two groups. Then, for purposes of comparison, *t* ratios were computed for the differences between the means of the two groups on each trait. These appear in Table 1. To simplify interpretation, profiles of the two groups, based on the mean scores, were plotted on the

Table 1

Means, Standard Deviations, Differences Between Means and *t* Ratios

Group	Test Variable	Mean	Standard Deviation	*t*
E	S	12.2	7.20	8.69**
S	S	19.8	8.61	
E	T	34.3	7.74	.60
S	T	34.8	7.89	
E	D	10.8	7.35	6.00**
S	D	16.1	8.49	
E	C	15.2	8.76	4.00**
S	C	19.5	10.48	
E	R	41.2	8.56	3.19**
S	R	38.0	9.68	
E	G	12.9	4.22	4.27**
S	G	10.8	4.64	
E	A	26.0	4.58	6.81**
S	A	21.8	6.33	
E	M	22.3	4.46	.23
S	M	22.4	4.48	
E	I	42.1	5.04	5.48**
S	I	38.5	6.93	
E	N	34.1	6.42	3.27**
S	N	31.5	7.77	
E	O	59.6	9.30	3.66**
S	O	55.3	11.70	
E	Ag	40.4	8.27	2.20*
S	Ag	38.3	9.24	
E	Co	78.0	14.00	6.54**
S	Co	67.0	16.30	

* Significant at the 5 per cent level.
** Significant at the 1 per cent level.

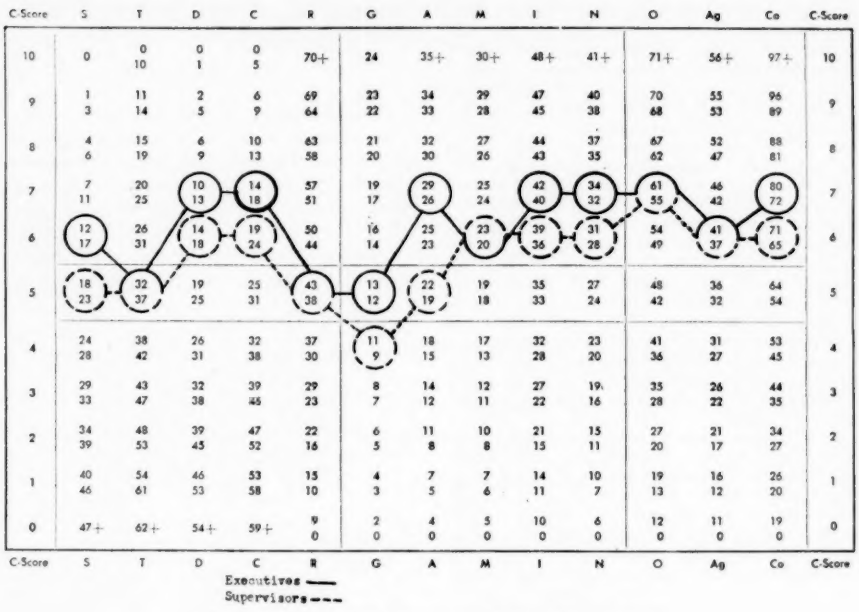

FIG. 1. A comparison between executives and supervisors.

background of the norms obtained for a population heavily weighted with college students, provided by the authors of the tests. These profiles appear in Figure 1.

The next procedure was to compute Pearson *r* correlations between test scores and ratings for each of the traits in the two groups. Since the ratings covered such a small range, it was necessary to correct the coefficients for coarse grouping. The correlations are presented in Table 2.

Assuming that the difference in level of position indicates to some extent a variable of supervisory success or excellence, scores on each trait were correlated with this executive-supervisor dichotomy using the point-biserial-*r* formula. The point-biserial correlations appear in Table 3.

Results

From Table 1 it can be seen that the executive averages significantly more: (1) sociable, (2) free from depression, (3) emotionally stable, (4) happy-go-lucky or impulsive, (5) active, (6) ascendant or socially bold, (7) self-confident or free from inferiority feelings, (8) calm and composed or free from nervousness, (9) objective, (10) agreeable, and (11) cooperative than the supervisor. Since the executive can be assumed to be more successful than the supervisor (if success is defined in terms of status, responsibility and income), it can also be assumed that he will possess a somewhat different pattern of traits than will the supervisor, if temperament contributes to his success.

It can be seen from the profiles (Fig. 1) that while these differences are evident, the more outstanding feature is the close similarity between the two profiles. The personality patterns might be said to follow the same general *trend*. It might also be noted that both of the groups are generally quite a bit above the population mean represented by college students. This deviation appears to be in the direction of "better adjustment."

The validity coefficients proved to be low in most cases. For executives the variables

Table 2

Pearson *r* Correlations Between Scores on Fourteen Test Variables and the Criterion of Job Performance Ratings

Group	Variable	Pearson *r**
E	S	**−.16**
S	S	−.03
E	T	−.01
S	T	+.03
E	D	−.09
S	D	−.13
E	C	−.04
S	C	**−.19**
E	R	.12
S	R	−.13
E	G	−.04
S	G	−.02
E	A	.07
S	A	.03
E	M	**.18**
S	M	.11
E	I	**.15**
S	I	.15
E	N	.02
S	N	**.18**
E	O	−.02
S	O	.16
E	Ag	.06
S	Ag	.10
E	Co	**.15**
S	Co	**.18**

Note: Correlations printed in bold-face type are significant at the 5 per cent level.

* For variables S, T, D and C a negative correlation indicates a positive relationship between the criterion and the traits of sociability, thinking introversion, freedom from depression and emotional stability.

which correlated significantly with the criterion of success were: (1) sociability, (2) freedom from inferiority feelings, (3) cooperativeness, and (4) masculinity. For supervisors the significant Pearson *r*'s were for variables: (1) emotional stability, (2) freedom from nervousness, (3) cooperativeness. Previous studies have indicated that the trait of cooperativeness as measured by the *Personnel Inventory* has most consistently shown a relationship with vocational success.

The interpretation of the validities obtained is that the more successful the executive, the more sociable, self-confident, cooperative and masculine he will tend to be (to an extent defined by the size of the correlation in each case). Greater degrees of emotional stability, composure, and cooperativeness will be expected of the more successful supervisor.

Because in some cases the regressions suggested some curvilinearity, it was considered advisable to test for this effect. The resultant chi squares indicated that in only one case was there significant departure from rectilinearity and this was in the case of factor *T* (thinking introversion) in the executive group, which probably means that too much or too little of this trait may be detrimental to success in this position.

In order to determine the effectiveness of scores on the traits correlating significantly with the criterion in terms of their contribution to selection of successful executives and supervisors, reference was made to the Taylor-Russell tables (**6**). It was found that for a selection ratio of .05, improvement in selection of "A" rated executives with use of scores on traits S, M, I and Co will be 10%, 11%, 9%, and 9% respectively, and of "A" rated

Table 3

Point Biserial Correlations Between Scores on Eleven Test Variables and the Executive-Supervisor Dichotomy

Variable	Point Biserial *r*
S	−.43
D	−.31
C	−.22
R	.17
G	.23
A	.36
I	.30
N	.18
O	.20
Ag	.12
Co	.34

Note: Factors S, D and C correlate negatively with the criterion because lower mean scores on these traits indicate greater amounts of sociability, freedom from depression and emotional stability, respectively.

supervisors improvement from use of scores on traits C, N and Co will be 18%, 17%, and 17%, respectively. Assuming normality of distribution, Taylor-Russell methods are applicable because each of these variables shows a rectilinear relationship with the criterion. A selection ratio of .05 was considered most appropriate since only a very small proportion of the industrial population is chosen for these higher-level jobs.

The point-biserial correlations are higher than are the Pearson *r*'s for the separate groups. They are, however, lower than they might be had the division of the dichotomy been more even than the necessary 208–143 split. The variables used are those in which significant differences were found between two group means since it is only here that a genuine correlation can be assumed to exist. It might be possible to establish cutting scores for those variables which show point-biserials larger than .20. On the basis of the cutting scores an applicant might be predicted to fall within one or the other of the two groups.

Certain limiting factors have exerted their influence upon the size of the obtained validity coefficients. The most significant of these limitations is selectiveness of the groups. This selection would restrict the variance on these trait scores and consequently would lower the validity coefficients from those which would have been obtained from a more heterogeneous population of job applicants. Another limiting factor is attenuation due to the unreliability of the criterion. If correction for attenuation were applied for errors in the criterion measure, the validity coefficients would probably be about 20% higher. Also, since there is no way at this time of determining the extent to which responses were falsified, this falsification may have served to lower the coefficients unless this source of variance is also correlated with the criterion.

What is needed is a set of norms for these tests from a large industrial population consisting not only of executives and supervisors, but covering the entire range of personnel. In order for these norms to be of much practical value, however, some means of overcoming possibilities of falsification must be devised. As Ruch (**5**) has suggested, each personnel department might perform a trial testing in which the executive or supervisor is asked first to respond to the inventory honestly and then again as he thinks the best executive or supervisor would respond. He must, of course, be assured that his score will in no way affect his status. In this way questions could be weighted according to the incidence of falsification on the second trial and thus a validating key could be devised. This may seem an unduly involved and expensive procedure but in view of the absence of selective devices for these all-important positions, it might well be worth the time and expense entailed.

Summary

On the basis of the investigation of 208 executives and 143 supervisors in a large chain grocery with regard to temperament traits as revealed by scores on the Guilford personality inventories, it was found that:

1. The executive averaged significantly more: (1) sociable, (2) free from depression, (3) emotionally stable, (4) happy-go-lucky, (5) active, (6) ascendant or socially bold, (7) self-confident, (8) calm and composed, (9) objective, (10) agreeable, and (11) cooperative than did the supervisor.

2. On the basis of a criterion of success defined as "job performance" ratings, the following traits contributed significantly to success of the executive: (1) sociability, (2) lack of inferiority feelings, (3) cooperativeness, and (4) masculinity. The traits contributing significantly to the success of the supervisor were: (1) emotional stability, (2) calmness and composure, and (3) cooperativeness.

3. Improvements in prediction of executive and supervisor success by means of the Taylor-Russell technique were, for the "A" rated executives for traits S, M, I and Co, 10%, 11%, 9%, and 9%, respectively, and for the "A" rated supervisors for traits C, N, and Co, 18%, 17% and 17%, respectively.

4. The obtained validity coefficients are lower than they might be because of the factors of restricted range of scores due to selectiveness of the groups, possibility of falsification of responses on the tests, and attenuation in the criterion.

5. Point-biserial correlations of test scores

and the executive-supervisor dichotomy as the criterion, reveal higher validity for many of the trait scores and indicate the possibility of prediction into one or the other category on the basis of a cutting score.

6. It is suggested that these results, while they agree with previous findings, be regarded as highly tentative until a great deal of further investigation confirms the stability of the obtained validities. It is also proposed that some attempt be made to determine the extent of falsification by weighting items according to the extent to which executives and supervisors are able to alter their scores in an experimental situation. There is great need for norms for these temperament trait scores based on an industrial population since it is generally accepted that success in executive and supervisory positions is largely dependent upon personality characteristics.

Received July 23, 1951.

References

1. Guilford, J. S. *The relative value of fourteen test variables for predicting success in executive and supervisory positions.* Unpublished Master's Thesis, University of Southern California Library, 1951.
2. Guilford, J. P. *An inventory of factors STDCR. Manual of directions and norms (revised).* Beverly Hills: Sheridan Supply Company.
3. Guilford, J. P., and Martin, H. G. *Inventory of factors GAMIN. Manual of directions and norms (first revision).* Beverly Hills: Sheridan Supply Company.
4. Guilford, J. P., and Martin, H. G. *The Guilford-Martin Personnel Inventory. Manual of directions and norms.* Beverly Hills: Sheridan Supply Company.
5. Ruch, F. L. *A technique for detecting attempts to fake performance on the self-inventory type of personality test. Studies in personality.* New York: McGraw-Hill Book Co., 1942.
6. Taylor, H. C., and Russell, J. T. The relationship of validity coefficients to the practical effectiveness of tests in selection. *J. appl. Psychol.*, 1939, **23**, 565–578.

Psychological Test Performance of Steel Industry Production Supervisors

Wesley A. Poe and Irwin A. Berg

Northwestern University

In addition to specific knowledge of the jobs they direct, supervisors in large industrial plants typically are expected to demonstrate ability in a number of other areas. Many of them, for example, serve as buffers for management when initial questions of collective bargaining agreements are raised. They are expected to apply intelligently information supplied by technical specialists, and they give instructions to the employees whom they supervise. Further, they must be intimately familiar with the flood of management control devices in the form of blueprints, charts and mimeographed directives. Such broad job demands would require that considerable care be exercised in the selection and training of supervisors.

As Lawshe (**8**) has noted, great emphasis has been placed upon the training of supervisors but relatively little has been done about their selection. Attempts to utilize psychological tests for selecting supervisors have met with varying success. Shuman (**12**) found that supervisors rated "excellent" in several plants could be identified by a battery of three tests. Harrell (**7**) found that the *Otis Test of Mental Ability* had low but useful discriminatory value for cotton mill supervisors who were rated as "satisfactory" by their superiors. In a study of 44 industries in each of which the two "best" and two "poorest" supervisors were given the *Adaptability Test*, Lawshe (**8**) found higher percentages of supervisors appearing in the "best" category as successively higher cut-off scores were used.

The present study seeks to assess the value of certain psychological tests for the purpose of identifying "good" production supervisors in a large midwestern steel-manufacturing plant. Present selection methods in this firm include tryouts on the job, interviews with management officials, and informal comments by foremen.

Procedure

There were 40 production supervisors at the department head or assistant department head level in the steel plant at the time of the present investigation. Seven of the 40 were unwilling or unable to participate in the study. Thus our data are derived from the performance of the 33 men who completed all tests in the battery described below. These men were the line supervisors, the "production bosses" who were directly in charge of fabricating the steel products sold by their company. All of them supervised the men who actually made the product, e.g., workers who ran punch presses, shapers, grinders, lathes, forge hammers, and the like. The types and scope of plant operations under the jurisdiction of the 33 supervisors were roughly the same, and their duties were reasonably comparable. Further, all of these men were full-time supervisors. Particular emphasis is placed upon the comparability of their responsibilities because some studies have ignored this important aspect. If research findings are to have value for other companies engaged in similar work, production foremen should not be grouped in a statistical potpourri along with supervisors of inspection, die-sinking, maintenance, and other departments. The present study seeks to avoid this pitfall.

A battery of ten tests was used in the study. The test administration was preceded by a portion of the *Otis S. A. Test* which was used for "warm-up" purposes. The tests used were: *Survey of Space Relations* (**4**); *California Short-Form Test of Mental Maturity* (**14**); *Adaptability Test* (**15**); *Test of Mechanical Comprehension* Form BB (**2**); *Survey of Mechanical Insight* (**11**); *The Bernreuter Personality Inventory* (**3**); *Guilford-Martin Personnel Inventory I* (**5**); *Inventory of Factors GAMIN* (**6**); *Strong's Vocational Interest Blank* (**13**); and Allport-Vernon's *A Study of Values*

(1). The entire battery required 7 hours to administer and was given in one day.

Each of the 33 supervisors was rated on his success as a supervisor by the plant superintendent and, independently, by the two assistant superintendents. Because the paired-comparison technique of rating (9) has purportedly considerably higher reliability than most rating methods, this approach was used in the form of the *Personnel Comparison System* (10). This procedure employs a booklet composed of slips of paper with two names written on each slip. The rater simply checks the preferred name on each slip. The number of times each individual is preferred is then tallied. Although there was some disagreement about supervisors in the middle positions, the raters were unanimous in their agreement on the top 10 and bottom 10 supervisors. These supervisors, representing those rated at the top 30 per cent and the bottom 30 per cent, were designated as the "high" and "low" criterion groups. After the criterion groups were identified, the test performance of the two groups was compared.

The *t* test was used to establish the significance of test score mean differences between the "high" and "low" rated groups of supervisors. *F* ratios were also computed as a check on the randomness and normality of the distributions. As scrutiny of the tables will disclose, the assumption of homogeneity of variance has been met in all but two cases.

Results

It is difficult to evaluate the effect on test performance of the mean age and education differences between the "high" and "low" groups of supervisors as shown in Table 1. While normal deterioration with age may be reflected in some of the test score differences, it seems unlikely that the original ratings should have been similarly influenced by the ages and educational levels of the men rated. Such an assumption would require that all three of the raters would have adopted similar stereotypes and worn identical "halos" since their agreement on which supervisors belonged in the top and bottom groups was perfect.

All of the supervisors in the present study were doing a satisfactory job and being paid

Table 1

Characteristics of the Subjects of the Criterion Groups

Characteristic	Group "High"	Group "Low"
N	10	10
Age:		
Mean	45.2	56.1
S.D.	4.8	5.7
Range	39–53	46–63
Education (in years):		
Mean	15.6	10.5
S.D.	1.2	2.5
Range	12–16	8–16
Rating score range	57–77	23–43

6 to 10 thousand dollars a year for doing it. The "high" and "low" groups represented men rated as performing in a clearly superior fashion, on the one hand, and men doing an ordinary but acceptable job on the other. Had any of them revealed obvious evidence of deterioration they would not have long remained in their positions. Further, since these men were carefully selected from a large group of employees of demonstrated capacity, it seems probable that any deficiencies among the supervisors studied were not readily apparent to the raters or, at least, not considered to be important. Thus it is believed reasonable to assume that the differences between the groups as shown in Table 1 did not markedly influence the ratings.

Whatever the possible influence of age and formal education, it seems clear that the mental ability tests used can discriminate usefully between the groups of supervisors studied. The differences between the total-score means on both mental ability tests (Table 2) were highly significant statistically. One part-score, *logical reasoning*, did not meet the assumption of homogeneity of variance. Another part-score, *spatial relations*, failed to distinguish acceptably between the two groups. By contrast, *The Survey of Space Relations* (Table 3) which presumably measures a similar ability in a different way did discriminate significantly between the "high" and "low" groups. The latter test, being longer and specifically validated for industrial groups, is probably the more appropriate instrument for the present investigation. Similarly, the

Mechanical Comprehension Test (Table 3) also distinguished significantly between the two groups of supervisors. These findings appear reasonable in that higher levels of mechanical ability among steel industry supervisors should be expected to contribute to a smoother running department from a production standpoint, at least where machine operations are concerned. Presumably, such men would tend to be rated higher on job performance. It is also probable that such men would be promoted sooner; hence the lower mean age of the "high" group.

Any competent observer of supervisors at work would agree that personality factors play an important part in supervisory success. The personality and interest tests used in the present study were included with the hope that some useful differences between the "high" and "low" groups would be found. The results, presented in Tables 4, 5, 6, 7, and 8, are quite disappointing in this respect.[1] The *Bernreuter* scales B2-S (Self-Sufficiency) and F2-S (Sociability) differentiated the groups at the 5 per cent level of confidence. Even these minimally significant differences are weakened since the two scales correlate .60 according to the manual (**3**). The *Strong Vocational Interest Blank* scores for *Aviator* and the *masculinity-femininity* scale also distinguished between the two supervisory groups at the 5 per cent level of significance. However, since 42 scales were scored on the *Vocational Interest Blank*, we should expect to obtain two scales differentiating at the 5 per cent level by chance alone. At any rate, further study is clearly indicated before the value of the *Strong* can be assessed for identifying successful supervisors in the steel industry. The *Personnel Inventory I, Inventory of Factors*

Table 2

Performance on the Mental Ability Tests

Test	Group "High"	Group "Low"	*t* Ratio	*F*
California Short-Form Test of Mental Maturity				
Total Mental Factors				
Mean	101.1	63.0	5.24**	4.04
Sigma	9.7	19.5		
Language Factors				
Mean	59.8	31.9	5.60**	3.22
Sigma	7.3	13.1		
Non-Language Factors				
Mean	41.3	29.6	4.30**	1.07
Sigma	5.9	5.7		
Spatial Relationships				
Mean	26.3	20.6	1.64	1.49
Sigma	5.0	4.1		
Logical Reasoning				
Mean	21.7	14.1	2.25*	39.89†
Sigma	1.6	10.1		
Numerical Reasoning				
Mean	18.9	9.4	4.86**	1.05
Sigma	3.8	3.7		
Vocabulary				
Mean	34.2	17.4	4.26**	2.10
Sigma	6.7	9.7		
Adaptability Test				
Mean	26.7	18.1	3.75**	4.41
Sigma	3.0	6.3		

* All differences so marked are significant at the 5 per cent level of confidence = 2.10.

** All differences so marked are significant at the 1 per cent level of confidence = 2.88.

† All differences so marked fail to meet the assumption of homogeneity of variance. Maximum *F* value of 4.5 for significance beyond 5 per cent level.

Table 3

Performance on the Mechanical Ability Tests

Test	Group "High"	Group "Low"	*t* Ratio	*F*
Mechanical Comprehension				
Mean	48.4	36.9	2.23*	2.43
S.D.	8.4	13.1		
Survey of Mechanical Insight				
Mean	31.0	16.3	4.34**	20.47†
S.D.	2.1	9.5		
Survey of Space Relations				
Mean	78.2	51.1	3.54**	1.58
S.D.	14.3	18.0		

* Significant at 5 per cent level.
** Significant at 1 per cent level.

[1] To save space and to reduce printing costs, Tables 4, 5, 6, 7, and 8 have been deposited with the American Documentation Institute. Order Document 3630 from American Documentation Institute, 1719 N Street, N.W., Washington, D. C., remitting $1.00 for photocopies (6×8 inches) readable without optical aid or $1.00 for microfilm (images 1 inch high on standard 35 mm. motion picture film).

GAMIN, and the *Study of Values* failed to yield useful differences between the "high" and "low" groups of supervisors.

Since both groups of supervisors were doing at least a satisfactory job, it is possible that no important, measurable differences in personality or interest distinguished one group from the other, despite the fact that the "high" group was rated as doing a much better job than the "low" group. On the other hand, the tests employed may not have accurately reflected any existing differences. One has the feeling, if opinion may be permitted, that real differences of interest and personality do exist between such groups. Perhaps other instruments may demonstrate the presence of any such supposed differences.

Summary

The job success of 33 production supervisors in a large steel-manufacturing concern was independently rated by the plant superintendent and two assistant superintendents using a paired-comparison rating technique. The 10 supervisors rated highest and the 10 rated lowest were given a 7-hour battery of 10 tests of intelligence, mechanical ability, personality, and interest. Differences in test scores between the "high" and "low" groups of supervisors were studied. It was found that the *California Short Form Test of Mental Maturity*, the *Survey of Space Relations*, and the *Adaptability Test* differentiated between the "high" and "low" supervisors at the 1 per cent level of confidence. The *Mechanical Comprehension Test*, two scales on the *Bernreuter Personality Inventory*, and two scales on the *Strong Vocational Interest Blank* distinguished between the two groups of supervisors at the 5 per cent level of significance. None of the other tests yielded significant differences.

Received September 17, 1951.

References

1. Allport, G. W., and Vernon, P. E. *Manual for the study of values.* Cambridge: Riverside Press, 1931.
2. Bennett, G. K. *Manual for the test of mechanical comprehension.* New York: Psychological Corporation.
3. Bernreuter, R. G. *Manual for the personality inventory.* Stanford University: Stanford University Press, 1935.
4. Case, H. W., and Ruch, F. *Manual for the survey of space relations.* Los Angeles: California Test Bureau.
5. Guilford, J. P., and Martin, H. G. *Manual for the personnel inventory I.* Los Angeles: Sheridan Supply Company.
6. Guilford, J. P., and Martin, H. G. *Manual for the inventory of factors GAMIN.* Los Angeles: Sheridan Supply Company.
7. Harrell, T. W. Testing cotton mill supervisors. *J. appl. Psychol.*, 1940, **24**, 31–35.
8. Lawshe, C. H. How can we pick better supervisors? *Personnel Psychol.*, 1949, **2**, 69–74.
9. Lawshe, C. H., Kephart, N. C., and McCormick, E. J. The paired comparison technique for rating performance of industrial employees. *J. appl. Psychol.*, 1949, **33**, 69–77.
10. Likert, R. *Morale and agency management.* Hartford: Life Insurance Sales Bureau, 1941.
11. Miller, D. R. *Manual for the survey of mechanical insight.* Los Angeles: California Test Bureau.
12. Shuman, J. T. The value of aptitude tests for supervisory workers in the aircraft engine and propeller industries. *J. appl. Psychol.*, 1945, **29**, 185–190.
13. Strong, E. K. *Manual for the vocational interest blank.* Stanford University: Stanford University Press.
14. Sullivan, E. T., Clark, W. W., and Tiegs, W. W. *Manual for the California short-form test of mental maturity, advanced form.* Los Angeles: California Test Bureau, 1947.
15. Tiffin, J., and Lawshe, C. H. *Preliminary manual for the adaptability test.* Chicago: Science Research Associates, 1943.

A Cluster Analysis of Office Operations *

Leon L. Thomas

Dunlap and Associates, Inc., Stamford, Connecticut

The identification of components of work is an integral step for the investigator in the area of Industrial Psychology. One phase in the construction of tests for selection and placement purposes is the determination of groups of similar operations. The development of a training program depends, in part, upon the isolation of components of work that tend to cut across job lines. Likewise, work components contributing to the make-up of a job family can serve as a foundation upon which a system for rating job incumbents can be built. Investigators in these areas are constantly seeking methods which will give them more information about the fundamental nature of the job families under consideration.

Research by Lawshe and others (**4**, **5**) suggests that factorial methods can be successfully applied to job analysis data. They have shown that factorial methods can be used to identify components when applied to job evaluation systems. Tryon (**6**) has developed a technique intended to accomplish results similiar to those of the factorial methods in a somewhat less mathematically rigorous manner. The Tryon technique seemed promising and a modification was developed for use with this particular problem.

This study, therefore, represents a new methodological attack upon the identification of components of work. Specifically, the purpose is to: (1) identify groups, or clusters, of similar elemental operations in a sample of office jobs; and (2) develop a modification of the Tryon Cluster Analysis technique for use with a large number of variables.

Procedure

The Check-List. Largely through the work of Culbertson (**1**) and Dudek (**2**), the *Job Description Check-List of Office Operations* has been constructed for use in the analysis of office jobs.[1] It consists of a series of 139 basic clerical operations. The check-list, for purposes of this study, was completed by having each job incumbent and his or her immediate supervisor check, independently, the duties performed on the job. A third party, usually the coordinator of research within the company, compared the two and identified any points of difference. A conference was then held with the incumbent and the supervisor in order to reach agreement.

The Sample. The sample of office jobs used in this study was obtained from five different companies. One was a foundry located in the South, another was an office filing equipment manufacturer in the East. A third contributor was a member of the automotive industry in the Midwest and two were steel mills, one located in the Midwest, the other in the East.

The Selection Criteria. Each of the five companies was asked to return completed check-lists for 25 key jobs. The key jobs were to be distributed throughout the entire range at present pay rates; were to "sample" the various areas of work being performed; and were not to be in dispute with regard to pay rates. From the 115 check-lists so obtained (the foundry was able to supply only 15), 112 were used in this study. Three were discarded because of insufficient data.

Limitations of the Sample. It is to be remembered that the results of this study are based upon this particular sample of office jobs. These companies are engaged in relatively different kinds of activity, and the jobs obtained from them seem to sample the various areas of office operations. It is possible, however, that a similar analysis made of office operations obtained from a different population might produce different results.

The Selection of Items. A frequency count was made of the number of times each check-list item was used in the sample of 112 jobs. The frequencies ranged from zero to 98, and had a median of 32. Seventy-nine items were found to have frequencies of twenty or more. These 79 items were the ones used in this analysis.

* This study is based upon a thesis submitted in partial fulfillment for the degree of Doctor of Philosophy, done under the direction of Professor C. H. Lawshe. The original data for this study can be found in the author's thesis on file in the Purdue University library.

[1] To reduce printing costs the *Job Description Check-List of Office Operations* has been deposited with the American Documentation Institute. Order Document 3267 from American Documentation Institute, 1719 N Street, N.W., Washington, D. C., remitting $1.00 for microfilm (images 1 inch high on standard 35 mm. motion picture film) or $1.20 for photocopies (6 × 8 inches) readable without optical aid.

The Correlation Matrix. A matrix of intercorrelation for the 79 selected items was computed. The standard formula, with certain adaptations, for the computation of ϕ was used[2] (3). The coefficients of correlation ranged from $-.25$ to $+.81$.

Construction of Tryon Curves. A curve was drawn according to the Tryon Method for each of the 79 selected items (6). The units along the ordinate were numbered from $-.25$ to $+.81$. The units along the abscissa were numbered from 1 through 79, the number of items in the analysis. The intercorrelations for a particular item were plotted by locating a point above each of the items along the abscissa at a distance corresponding to the size of the correlation between the item being plotted and each of the other 78. For the correlation of each item with itself, the value of the highest intercorrelation of that item with any of the others was used. The points on each of these 79 plots were joined producing a curve indicating the manner in which a particular item correlated with the other 78.

$$^{2}\phi_{ij} = \frac{ad - bc}{\sqrt{(a+b)(a+c)(c+d)(b+d)}}$$

where a = the number of jobs on which both item i and item j were checked; b = the number of jobs on which item i was checked and item j was not checked; c = the number of jobs on which item j was checked and item i was not checked; and d = the number of jobs on which neither item i nor item j was checked.

In order for distances in the region where ϕ approached $+$ or -1 to be equal to those in the region where ϕ approached zero, the values along the ordinate were transformed to Fisher's z'. Also, each curve was drawn on a separate $40'' \times 24''$ sheet of tracing paper so that the curves could be superimposed for subsequent analysis.

Curve Inspection. In a Tryon Analysis, groups or clusters of items are identified by isolating groups of congruous curves. In order to accomplish this with 79 separate curves, it was necessary to establish a starting point, or preliminary sets of items. First, a table was drawn up listing each of the selected items and the three items most highly correlated with them. From this table it was possible to note groups of from three to five items which exhibited a high degree of intercorrelation. Sixteen tentative clusters were identified, accounting for 40 of the 79 curves.

A viewing box was constructed, $40'' \times 24''$ in size with a translucent plastic top lighted by seven 60 w. lumiline bulbs. This made it possible to superimpose the curves for inspection.

First, the curves for the preliminary sets of items were inspected for congruity. On the basis of inspection some of the sets were combined, reducing the number of sets to seven, accounting for 40 of the items. Then each of the remaining curves was compared with each

Table 1

Distribution of Cluster Item and Residual Item Correlations After the First Refinement

	I		II		III		IV		V		VI		VII		VIII	
Corr.	C	R	C	R	C	R	C	R	C	R	C	R	C	R	C	R
.75–.79	1				1						1					
.70–.74	1		1				1				1					
.65–.69	3		3		1		2		3		1		1		1	
.60–.64	2		5		1		1						1			
.55–.59			3		2		2						5		1	
.50–.54			3		1		1		1		1	1	1			
.45–.49	1				3	1		1			1		2	1		1
.40–.44				3			1	2	1						1	
.35–.39						2		1		1				1		
.30–.34				2		4		4		1		1		2		1
.25–.29		1				3		3		3		2		3		2
.20–.24				3		2		3		1				1		1
.15–.19		2		2		2		2		3		5		5		
.10–.14		2		2						2		1		1		4
.05–.09		4		4		2				3		4				
.00–.04		3								2		1		1		5
−.00–.04		2										1		1		1
−.05–.09		2														1

C = Cluster items.
R = Residual items.

Table 2

Distribution of Cluster Item and Residual Item Correlations After the Second Refinement

	I		II		III		IV		V		VI		VII		VIII	
Corr.	C	R	C	R	C	R	C	R	C	R	C	R	C	R	C	R
.75–.79	1				1											
.70–.74	1		1				1				2					
.65–.69	3		3		2		1		3		1				1	
.60–.64	1		5		1		3						2			
.55–.59			3		1		2				1		4			
.50–.54			3		2				1		1		5		3	
.45–.49	1				2		2				1					
.40–.44				2	1		2		1							
.35–.39				1		1		1								
.30–.34						2		3		1		2				1
.25–.29		1				3		2		2		1		1		1
.20–.24						1		2						1		1
.15–.19		1		1				1				2		5		
.10–.14		1		2		1				1		2		1		
.05–.09		2		3		1				3		1				3
.00–.04		1								2		1				1
−.00–.04		1												1		2
−.05–.09		2														

C = Cluster items.
R = Residual items.

of the seven sets and 20 of them added to one or the other of the sets. Finally, the nineteen remaining curves were compared with one another and an eighth set emerged consisting of three curves. These eight sets of curves, identified by inspection, accounted for 63 of the 79 selected items. Sixteen of the curves exhibited no apparent congruity with any of the eight identified sets or with one another.

First Refinement. The correlation between each of the 63 items, which had been placed in one of eight clusters by inspection, and the cluster in which it appeared was computed.[3] Similarly, the correlation between each of the 16 residual items and each of the eight clusters was computed. Table 1 shows the distribution of correlations between the 63 cluster items and the cluster in which they appear and the distribution of the correlations between the 16 residual items and each of the eight clusters.

Second Refinement. It can be seen from Table 1 that seven of the residual items correlate as highly or more highly with one or the other of the eight clusters than some of the items appearing in them originally. It was decided to include these seven items in the clusters. One item was added to cluster III, three items to cluster IV, one item to cluster VI, one to cluster VII, and one item to cluster VIII.

After these additions, again the correlations between each of the 70 cluster items and the cluster in which it appeared were computed. Likewise, the correlations between each of the nine residual items and each of the eight clusters were computed. Table 2 shows the distribution of these correlations after the second refinement.

It can be seen from Table 2 that none of the correlations between the residual items and the

[3] The author is indebted to Mr. P. C. Baker of the Division of Educational Reference, Purdue University, for the development of this statistic.

$$r_{sI} = \frac{\Sigma r_{sIj}}{\sqrt{n_I + 2\Sigma r_{Iij}}}$$

where r_{sI} = the correlation of item s with cluster I; Σr_{sIj} = the sum of the correlations of item s with the items composing cluster I; Σr_{Iij} = the sum of the intercorrelations of all the items composing cluster I; n_I = the number of items composing cluster I.

Table 3

Matrix of Cluster Intercorrelation

Cluster	I	II	III	IV	V	VI	VII
II	−.011						
III	−.042	.589					
IV	−.050	.497	.607				
V	.263	.428	.426	.232			
VI	.186	.135	.158	.283	.173		
VII	.131	.392	.398	.185	.363	.384	
VIII	.097	.641	.430	.285	.387	.195	.308

eight clusters overlapped with those between the cluster items and the clusters in which they appeared. It was decided, therefore, to stop the addition of items after the second refinement and let the nine remaining items remain as residual or "unique" items.

Cluster Intercorrelations. Table 3 shows the intercorrelations of the eight identified clusters.[4] The residual items showed little tendency to intercorrelate. Their intercorrelations ranged from −.14 to +.34, with a mean of .14.

Results

Eight clusters, or components of work, constituting office operations in a sample of 112 office jobs were identified. These clusters of items may be characterized by the type of operations found within them.

Cluster I (*Typing*). It seems probable that the operations within cluster I group themselves because of the equipment used in the performance of the operations. It consists of such items as: "Typewrites straight copy from correct forms," "Prepares duplicating machine 'stencils' and/or 'master copies' using a typewriter," and "Typewrites straight copy from rough draft." All the items contained in cluster I pertain to the use of the typewriter in various ways, addressing, making out forms, checks, etc.

Cluster II (*Listing and Compilation*). Cluster II may be described as a cluster of listing and compilation activities. The items isolated in cluster II pertain to copying, checking, listing, and recording activity. It contains such items as: "Compiles lists of numerical, verbal, or other descriptive data," "Prepares routine lists of specific items, numerical and/or verbal, according to designated system by longhand," and "Copies desired data from one form or record into the proper place on another form or record by longhand."

Cluster III (*Communication*). Cluster III may be thought of as information gathering and dissemination activity. The items contained in this cluster deal with placing and answering telephone calls, sending, routing and relaying forms, materials, and information to proper persons or departments and receiving requests, instructions or information.

Cluster IV (*Planning and Supervision*). Planning and supervisory operations characterize cluster IV. Isolated in this cluster were items such as: "Plans and/or coordinates the assignment or execution of duties performed by other individuals," "Reviews and analyzes work of others, calling attention to use of incorrect procedures or methods," and "Composes correspondence requiring specific knowledge of methods, procedures, policies, or other information."

Cluster V (*Filing*). Cluster V can be described as a cluster of filing operations. Operations found in this cluster pertain to the classification, coding, positioning, location, and transference of materials in files.

Cluster VI (*Stock Handling*). The phrase "stock handling" is appropriate for describing cluster VI. Items pertaining to wrapping and packaging of articles, removing stock, merchandise, or other items from special places or containers, and marking items with identification and postal information are found in it.

Cluster VII (*Routine Clerical*). Cluster VII is more difficult to describe. The operations in cluster VII seem to pertain to routine, low-level, clerical activity. It contains items such as: "Counts various items," "Assembles various forms," "Fastens together various forms, records or other items," "Cuts or tears apart special forms," and similar activities.

Cluster VIII (*Calculation*). As can be seen from Table 3, cluster VIII is somewhat highly related to cluster II. This degree of relationship can be explained, at least in part, when the items in cluster VIII are compared with those of cluster II. Whereas cluster II contains several items pertaining to the copying, compilation, and listing of numerical data, cluster VIII contains items dealing with the mathematical manipulation of numerical data. It seems reasonable, therefore, that cluster VIII be thought of as a "dependent" cluster with regard to cluster II, hence, highly associated with it.

Residual Items. Nine of the original 79 selected items remained as residual items

[4] The author is indebted to Mr. P. C. Baker of the Division of Educational Reference, Purdue University, for the development of this statistic.

$$r_{I,II} = \frac{\Sigma r_{IiIIj}}{\sigma_I \sigma_{II}}$$

where $r_{I,II}$ = the correlation between cluster I and cluster II; Σr_{IiIIj} = the sum of the intercorrelations of the items composing cluster I with those composing cluster II; $\sigma_I = \sqrt{n_I + 2\Sigma r_{Iij}}$; $\sigma_{II} = \sqrt{n_{II} + 2\Sigma r_{IIij}}$.

after the second refinement. The correlations between these items and the eight clusters were not high enough for them to be included in one or the other of the clusters, and their intercorrelations were small. It is tenable, therefore, to assume that these residual items are concerned with "unique" operations.

There are two possible explanations for the "uniqueness" of the residual items. First, it would be possible for a residual item to be associated with items that have been placed in the identified clusters, but, due to the ambiguity of the item, associated in so random a fashion that it could not be placed in any one cluster. Second, it would be possible for a residual item to be a "specific" operation requiring 100 per cent of an incumbent's time, or at least so much of an incumbent's time that other operations have little opportunity to become associated with it. An examination of the residual items lends further support for the latter explanation. The residual items pertain to such activities as: "Gives out various materials, forms or other items upon request of customers, clients or other individuals," "Receives merchandise, materials, forms, telegrams, mail, or other items which are being delivered, or being returned for refund, replacement, or other purposes," and "Sends telegrams."

Unselected Items. The items used in this analysis were selected on the basis of their frequency of occurrence in a particular sample of office jobs. Only the 79 items that were checked at least 20 times were used, leaving 60 items unselected. The addition of unselected items could affect the analysis in three ways. First, additional items could correlate highly enough with one or the other of the previously identified clusters to be included in them. Second, additional items could show up as "unique" operations. Third, it would be possible for additional clusters to emerge with the addition of items in the analysis. For instance, none of the operations concerned with stenographic or automatic machine activity had sufficient frequencies to be selected for the study. It is therefore likely that had these items been used often enough to have been included in the analysis, a ninth, or stenographic cluster, and a tenth, or automatic machine operation cluster, would have been identified.

Summary

1. The purpose of this study was a) to identify clusters of similar elemental operations in a sample of office jobs, and b) to develop a modification of the Tryon Cluster Analysis technique for use with a large number of variables.

2. Using completed *Job Description Check-Lists of Office Operations* from 112 office jobs, the intercorrelations for 79 selected check-list items were computed.

3. A modification of the Tryon Cluster Analysis technique was developed and used to identify eight clusters of office operations which may be thought of as the work components of a sample of office jobs.

4. The clusters, or components of work, identified were: a) Typing, b) Listing and Compilation, c) Communication, d) Planning and Supervision, e) Filing, f) Stock Handling, g) Routine Clerical Operations, and h) Calculation.

5. A procedural outline for performing a cluster analysis with a large number of variables was written.[5]

Received August 1, 1951.

References

1. Culbertson, A. L. *The adequacy of an operational check-list for the general description of office jobs.* Unpublished M.S. thesis, Purdue University, 1947.
2. Dudek, E. E. *An operational approach to the evaluation of office jobs.* Unpublished Ph.D. thesis, Purdue University, 1948.
3. Guilford, J. P. *Fundamental statistics in psychology and education.* Second edition, New York: McGraw-Hill, 1950.
4. Lawshe, C. H., and Satter, G. A. Studies in job evaluation. 1. Factor analysis of point ratings for hourly-paid jobs in three industrial plants. *J. appl. Psychol.*, 1944, **28**, 189–198.
5. Lawshe, C. H., Dudek, E. E., and Wilson, R. F. Studies in job evaluation. 7. A factor analysis of two point rating methods of job evaluation. *J. appl. Psychol.*, 1948, **32**, 118–129.
6. Tryon, R. C. *Cluster analysis.* Berkeley, Calif.: Associated Students Store, University of California, 1930. (Lithoprinted.)

[5] To reduce printing costs *A Cluster Analysis of Office Operations* has been deposited with the American Documentation Institute. Order Document 3627 from American Documentation Institute, 1719 N Street, N.W., Washington, D. C., remitting $1.00 for photocopies (6 × 8 inches) readable without optical aid or $1.00 for microfilm (images 1 inch high on standard 35 mm. motion picture film).

A Factor Analysis of a Salary Job Evaluation Plan *

Allen H. Howard and Howard G. Schutz

Department of Psychology and Education, Illinois Institute of Technology

The purpose of this study is to analyze factorially the operation of a point system of job evaluation as used in rating salary or office-type jobs in a manufacturing plant, and to determine the basic factors involved in the functioning of the system and their significance in determining labor grade level. Previous similar or related studies by Lawshe and associates (**2**, **3**, **4**, **5**, **6**), Rogers (**7**), and others have indicated consistently that regardless of the type of job evaluation plan used, the number of elements rated, or other specific features of the installation, job grade level seems to be determined by the operation of a very limited number of factors, one of these often accounting for 95 per cent or more of the labor grade variance. These studies have demonstrated that in many job evaluation plans the component rating scales do not provide independent measurements, and further, that certain of the variables are not significantly related to the combined job rating.

* Revision of a paper presented at the April 27–28, 1951, Meeting of the Midwestern Psychological Association. The authors wish to acknowledge their indebtedness to Dr. W. A. Kerr for providing in his class in advanced statistical techniques the opportunity to do preliminary work on this analysis and to Dr. P. S. Shurrager for his critical reading of the manuscript.

Data and Procedure

The present study is based on data obtained in a large electrical manufacturing plant where grading of non-supervisory, office-type jobs by means of a grading plan similar to the NEMA salary plan has been going on for a dozen or more years. Job rating is performed with the cooperation of line supervision by carefully trained analysts, whose work is rigorously checked by their own supervisors.

The essentials of the grading plan are presented in Table 1 with the 12 job attributes classified in terms of the a priori *Skill*, *Responsibility*, and *Job Conditions* attribute categories established by the authors of the plan. Of the 11 scales having more than two divisions four (Nos. 1, 2, 3, and 12) are divided arithmetically; the remaining seven are divided irregularly with larger steps at the high ends.

The general characteristics of most of the attributes are suggested by their respective titles, but several require additional comment. Attribute 8, *Responsibility for Service Relations*, is concerned essentially with responsibility for (a) relations with customers and (b) the security of confidential personnel information. Attribute 9, *Responsibility for Work of Others*, "relates to non-supervisory responsibility for instructing other employees, assigning work to them, coordinating their efforts, and maintaining the flow of work within a group." Attribute 10, *Business Relations*, is "a measure of the extent to which the job necessitates meeting,

Table 1

Summary of Job Evaluation Plan Showing Total Point Range and Number of Scale Divisions for Each Attribute

	Attribute	Total Point Range	Number of Scale Divisions
Skill	1. Education	58	5
	2. Experience	110	6
	3. Analytical Ability	45	4
	4. Judgment	85	4
	5. Ingenuity	20	2
	6. Expression	35	4
Responsibility	7. Responsibility for Monetary Loss	50	4
	8. Responsibility for Service Relations	50	4
	9. Responsibility for Work of Others	15	3
	10. Business Relations	40	4
	11. Supervision Received	35	4
Job Conditions	12. Working Conditions and Physical Demand	30	4

dealing with, and influencing other people." Attribute 5, *Ingenuity*, is defined as inventive ability "involving the origination of new designs, processes, manufacturing or business methods, or adapting of existing designs, processes, manufacturing or business methods, to new or changed situations." *Ingenuity* has been a little used scale in the evaluation of the vast majority of office-type jobs in this company. In fact only 11 of the sample of 200 evaluations used in the present study received the higher of the two possible ratings of this attribute. Since under these conditions any correlation of *Ingenuity* ratings with other attribute ratings would be of negligible significance, this attribute was not included in the computations to be described.

Data consisted of 200 evaluations representing a wide range of jobs at all grade levels, including routine clerical occupations as well as more highly specialized jobs in such areas as purchasing, accounting, personnel, drafting, inspection, and production control.

Results

Eleven attribute ratings and job grade were intercorrelated by the product-moment method and the resulting matrix, which is presented in Table 2, factor analyzed by Thurstone's centroid technique. After the extraction of one factor several criteria indicated that additional factors would probably be the result of chance variance only. An attempted extraction of a second factor gave a meaningless result after as well as before rotation.

As indicated in Table 3, the solitary group factor determined by the factor analysis is common in some degree to all eleven of the attributes included in the study. It accounts for a major part of the variance in eight of them (Nos. 1 through 7, 10, and 11) and for approximately 99 per cent of the variance in job grade level. The magnitude of the latter relationship is a function of the relatively heavy weights assigned in the grading plan to these attributes, and of the greater variability of these attributes as compared with the others.

Those attributes heavily saturated with this factor represent elements which should be expected to vary directly with the general skill and ability required by the job. The factor, then, is similar to the primary "Skill Demands" factor of Lawshe and others and has been so designated. Three of the attributes, Nos. 8, 9, and 12, are only slightly weighted with the "Skill Demands" factor. In the correlation matrix these attributes show relatively little relation to each other, to job

Table 2

Intercorrelations of Attribute Ratings and Job Grade

Attribute	2. Experience	3. Analytical Ability	4. Judgment	6. Expression	7. Monetary Loss	8. Service Relations	9. Work of Others	10. Business Relations	11. Supervision Received	12. Working Conditions and Physical Demand	Job Grade
1. Education	.76	.78	.74	.71	.55	.16	.32	.68	.69	−.12	.86
2. Experience		.85	.81	.72	.69	.17	.40	.77	.76	−.12	.94
3. Analytical Ability			.81	.72	.67	.15	.29	.78	.74	−.19	.92
4. Judgment				.80	.69	.09	.35	.81	.88	−.17	.93
6. Expression					.62	.18	.31	.82	.81	−.35	.86
7. Monetary Loss						.14	.29	.60	.63	−.17	.79
8. Service Relations							−.04	.17	.12	−.26	.23
9. Work of Others								.24	.36	−.04	.40
10. Business Relations									.82	−.28	.88
11. Supervision Received										−.21	.88
12. Working Conditions and Physical Demand											−.12

Table 3

Factor Loading for Each Attribute and Job Grade

	Attribute	Factor Loading
Skill	1. Education	.82
	2. Experience	.90
	3. Analytical Ability	.89
	4. Judgment	.91
	6. Expression	.88
Responsibility	7. Responsibility for Monetary Loss	.75
	8. Responsibility for Service Relations	.21
	9. Responsibility for Work of Others	.38
	10. Business Relations	.88
	11. Supervision Received	.88
Job Conditions	12. Working Conditions and Physical Demand	−.27
	Job Grade	.99+

grade, or to the other eight attributes, and it reasonably can be assumed that they function as relatively independent and quite specific variables, although of little significance in determining job grade. If each one were represented in the matrix more than once, they should come out in the analysis (if based on a sufficiently large sample of evaluations) as a series of additional group factors.

As stated previously, Attribute 8, *Responsibility for Service Relations*, is concerned essentially with responsibility for (a) relations with customers and (b) the security of confidential personnel information. That these two aspects of Attribute 8 may vary independently is perhaps a partial explanation of the paucity of correlation between this attribute, the others, and job grade. However, it is reasonable to believe that while the two aspects of *Responsibility for Service Relations* may vary to some extent independently of each other, they also vary independently of the other attributes; that is to say, they represent real and objective job elements, which can be judged to a considerable extent as matters of fact without relation to the "Skill" and other variables.

A study of the functioning of Attribute 9, *Responsibility for Work of Others*, suggests that this attribute need not be expected to show a very close relation to the general level of skill or to any specific attribute, because practically all jobs may carry this responsibility in any degree. Thus a job demanding only little more than the minimum of skill and ability may involve partial responsibility for directing the efforts of a helper, or even of coordinating the efforts of a group of coworkers; as, for example, a typist required to coordinate the flow of work within a group of which she is a member.

Like *Responsibility for Service Relations* Attribute 12, *Working Conditions and Physical Demand*, is made up of what are probably two partially independent variables, (a) working conditions and (b) physical demand. Here again the independent variation of these two elements may serve to reduce the correlation between this attribute, the others, and job grade. The negative correlation between the *Working Conditions and Physical Demand* attribute and the other elements in the matrix accounts for its negative loading with the "Skill Demands" factor. In other words, the negative loading is the result of a tendency for those jobs representing the higher skill levels to require lesser amounts of visual demand and physical effort and to be performed in less objectionable surroundings than jobs at the lower skill levels. A similar relationship has been reported by other investigators (**1**, **3**, **4**).

Summary

The foregoing analysis suggests that as a measuring instrument the job evaluation plan functions in a manner quite different from that probably intended by its authors. A single factor corresponding to Lawshe's "Skill Demands" accounts for a major portion of the variance in eight of the attributes and for approximately 99 per cent of the variance in job grade. The remaining attributes, although of little significance in accounting for variance in job grade, seem to represent job elements which function independently of the "Skill" factor. Thus many of the variables in the plan are not being evaluated independently, and the arbitrary weights which have been assigned to others do not give any realistic indication as to the degree of their association with job grade.

Received August 19, 1951.

References

1. Chesler, D. J. Reliability and comparability of different job evaluation systems. *J. appl. Psychol.*, 1948, **32**, 465–475.
2. Lawshe, C. H., Jr., and Satter, G. A. Studies in job evaluation. 1. Factor analysis of point ratings for hourly-paid jobs in three industrial plants. *J. appl. Psychol.*, 1944, **28**, 189–198.
3. Lawshe, C. H., Jr., and Maleski, A. A. Studies in job evaluation. 3. An analysis of point ratings for salary-paid jobs in an industrial plant. *J. appl. Psychol.*, 1946, **30**, 117–128.
4. Lawshe, C. H., Jr., and Alessi, S. L. Studies in job evaluation. 4. Analysis of another point rating scale for hourly-paid jobs and the adequacy of an abbreviated scale. *J. appl. Psychol.*, 1946, **30**, 310–319.
5. Lawshe, C. H., Jr., and Wilson, R. F. Studies in job evaluation. 5. An analysis of the factor comparison system as it functions in a paper mill. *J. appl. Psychol.*, 1946, **30**, 426–434.
6. Lawshe, C. H., Jr., Dudek, E. E., and Wilson, R. F. Studies in job evaluation. 7. A factor analysis of two point rating methods of job evaluation. *J. appl. Psychol.*, 1948, **32**, 118–129.
7. Rogers, R. C. Analysis of two point-rating job evaluation plans. *J. appl. Psychol.*, 1946, **30**, 579–585.

A Factor Analysis of Employee Attitudes *

Robert J. Schreiber, Robert G. Smith, Jr., and Thomas W. Harrell

University of Illinois

There is a lack of factorial studies on employee attitudes such as have been performed in the field of social attitudes (**1**, **3**, **4**, **5**, **8**, **11**, **14**). Goodwin Watson (**15**, p. 360) has presented an extensive list of attitude factors gained from employee verbalizations, and Strong (**12**, pp. 529–530) has presented a list of motivations that have been compiled from the existing literature.

What are the basic factors in employee attitudes? Ghiselli and Brown (**7**, p. 435) suggest that they are attitudes toward management and the working situation, and beliefs as to what constitutes company policy. While it is possible to hypothesize various factors, the hypotheses will, naturally, vary somewhat with the individual who hypothesizes. One method of solution is factor analysis.

Hull and Kolstad (**10**, p. 349) attempt to name factors by applying social attitude factors to employee attitudes. The validity of this remains in doubt until we have definitely ascertained that social attitudes and employee attitudes are interrelated.

The Attitude Variables and the Sample

During the academic year 1947–1948, the attitudes of the non-academic employees of the University of Illinois were surveyed by the interview questionnaire method. The survey covered a random sample of over 12% of the non-academic employees at the four campuses: the Galesburg undergraduate campus, the Navy Pier undergraduate campus in Chicago, the Urbana campus, and the University of Illinois Medical Center in Chicago. The results are reported in (**6**) and (**9**).

Five broad areas of job satisfaction were included in the survey, one of the purposes of which was to evaluate the effectiveness of various personnel policies in promoting job satisfaction. These a priori classifications were: (1) adjustment to job; (2) supervision; (3) participation-expression; (4) working conditions; and (5) incentives.

Since supervisory personnel are often identified with management by the rank-and-file worker, and since differences in attitudes of supervisory and non-supervisory employees have been found (**10**, p. 358), only the latter group are considered in this study. The size of the sample is 379.

Intercorrelations of the Variables

The responses were given weights of one or zero. In general, ones represent positive responses, while zeros are negative. There was a small percentage of *No Answer* responses, in no case representing more than 6% of responses to any one question. The *No Answer* responses were considered with regard to their meaning for each question, and weights were decided by two judges. These weights are indicated in the questionnaire.

The weights were so arranged as to dichotomize the responses as close to the median as possible and were inspected by the judges to insure that clearly negative responses were not classified as positive, or vice versa.

Computing diagrams for tetrachoric correlations (**2**) were used to determine the coefficients. At least two determinations, and if possible, more, were made, and the mean value used in an effort to increase the reliability of the correlation.

Over half of the inter-item correlations lie between .10 and .30. While these are not high, they are useful for indicating trends.

The mean intercorrelations between the a priori classes of items are shown in Table 1. It should be noted that the standard error of a tetrachoric r of 0.00 where $N = 379$ is .07. Hence, any r larger than .18 is significantly different from 0.00 at the 5% limit.

It is noteworthy that the correlations

* The present paper is a summary and an extension of the Master's Thesis by Schreiber, now with Dunlap and Associates, Inc., New York, N. Y. The thesis was supervised by Harrell. Smith had charge of the rotation of axes and assisted in writing an early draft of this paper. The University of Illinois Graduate Research Board gave financial support for statistical assistance in carrying out the analysis.

Table 1

Mean Intercorrelations Between Groups of Items

	I*	II*	III*	IV*	V*
I*					
II*	.22				
III*	.21	.30			
IV*	.23	.24	.25		
V*	.10	.13	.02	.16	

* I. Adjustment to Job: Items 1–4; II. Supervision: Items 5–10; III. Participation-Expression: Items 11–14; IV. Working Conditions and Facilities: Items 15–17; and V. Incentives: Items 18–28.

between Area V, Incentives, and the other areas are lower than the correlations between any of the remaining areas. Inspection of the item intercorrelations indicates that this is due to items 24–27, items which are more informational than attitudinal. This precludes many conclusions that might be drawn about the relationship of Area V with other areas. From the evaluation of the intercorrelations of the other broad areas, and consideration of the number of positive and negative r's affecting the mean, it is possible to draw some conclusions:

1. There is a small negative relationship between satisfaction with participation-expression and knowledge of University policies.
2. Satisfaction with supervision and satisfaction with participation-expression are positively related.
3. The quality of supervision is positively related to satisfaction on the job.

From an examination of some of the higher correlations between individual items ($r > .35$), it is possible to conclude that:

1. There is an indication that women employed by the University more often than men say that they know what is expected of them on the job when they begin work here.
2. Persons with greater amounts of educa-

Table 2

Loadings on Factor I, Job Satisfaction

Loading	Item
71	5. Do you feel free to speak to your foreman or supervisor and ask him for help or an explanation?
66	9. When you have done a good job are you ever given credit or complimented?
64	11. How much dissatisfaction do you think there is among workers in your department?
63	16. How friendly and cooperative is the feeling between the people with whom you work?
60	3. To what extent did you know what was expected of you on the job when you first began to work here?
54	18. Do you think you will be able to hold your present job as long as you want to?
53	13. Do you feel free to express your "gripes" and dissatisfactions to higher authority?
52	8. How much freedom do you feel that your supervisor or foreman allows you in doing your job?
52	14. Are you always informed when there are changes in rules which affect you?
51	12. What is the most important grievance that employees have? (lack of grievances)
47	17. Satisfaction with tools and equipment.
47	23. Do you feel that good, hard work will pay off in larger opportunity or pay?
46	4. Did the instructions first given to you as to your job and duties help much in making the "breaking in" period easier?
44	6. Is your supervisor always able to assist or explain details of your job?
42	7. How well does your supervisor know University policies, rules, or procedures?
38	1. How much were you able to use your past experience or training when first hired by the University?
38	2. To what extent did you know what was expected of you on the job when you first began to work here?
38	10. How much have you to say about the way your work should be done? (May you offer suggestions or change a way of performing a task?)
38	15. How do your general working conditions compare with other places in the city?
37	20. How much does your particular job allow growth or development of your knowledge or skill?
38	22. To what extent, if any, are opportunities provided for promotion and advancement?

Table 3

Loadings on Factor II, Knowledge of Employee Benefits

Loading	Item
71	26. Under the University policy, what provisions are made for your beneficiary in case of your death?
70	25. How much time off with pay is allowed each year for extended sick leave?
68	34. Member of retirement system.
67	27. What is the University plan for life income following retirement from actual service?
63	31. Years of service.
58	28. What do you think of the University's attitude toward the unions?
50	24. How much time off with pay is allowed each year for sick leave?
−48	33. Marital status (single).
−66	30. Sex (male).
−69	29. Age (below 30).
54	32. Educational background (not over 8 years).

tion tend to feel that their jobs will allow them more growth and development of their skills than do persons with less education.

3. Persons who are members of the retirement system more often than not see little opportunity for advancement in their jobs.

4. There is a moderate correlation between thinking that there is opportunity for advancement and thinking that hard work will pay dividends.

5. Another moderate correlation exists between having freedom in doing the job and freedom in expressing dissatisfactions.

6. Having satisfaction with tools and equipment is related to lack of grievances.

Factor Analysis

The intercorrelations between variables were factor-analyzed by the Complete Centroid Method of Thurstone (**13**, p. 161H).[1] The unknown communalities were estimated by inserting the highest coefficient in each column into the diagonal cell of that column. The communalities were re-estimated after the extraction of each factor.

[1] The final questionnaire and the complete tables of intercorrelations of the variables, together with complete factor analysis tables, have been deposited with the American Documentation Institute. These are: Table 4, Table of Intercorrelations; Table 5, Unrotated Factor Matrix; Table 6, Transformation Matrix; Table 7, Directional Cosines of the Factors; Table 8, Oblique Factor Matrix V. Order Document No. 3430 from American Documentation Institute, 1719 N Street, N.W., Washington 6, D. C., remitting $1.00 for microfilm (images 1 inch high on standard 35 mm. motion picture film) or $1.35 for photocopies (6 × 8 inches) readable without optical aid.

Eight sets of factor loadings were rotated blindly, once orthogonally, and once obliquely, to simple structure. After rotation it appeared that eight factors were undoubtedly too many, since clear interpretation could be given to only two.

Interpretation of the Factors

It was decided to use only variables with loadings of .40 or higher. However, since the factor loadings are actually significant only to the first digit, this criterion has been used with that in mind.

The first factor has been called "Job Satisfaction." The variables it loads .35 or higher are listed in Table 2.

The variables which this factor loads significantly are those which both previous studies and a priori considerations would indicate made for satisfaction with the job if a sufficient number of these are answered favorably. The importance of good supervision is underscored, since about half of the variables directly concern supervision. There are, however, enough variables dealing with other aspects of job satisfaction to warrant not calling this factor "Supervision." This is a factor of wide scope, since the only variables not significantly loaded with this factor are those dealing with knowledge of personnel policies and personal information items.

There is one variable, which, while dealing with an aspect of satisfaction with work, has a factor loading greater than zero, but not high enough to be used in interpretation of

Factor I. This is Item 19, which deals with satisfaction with pay. This variable has a low communality. Some other items which were used in interpretation of the factor also have low communalities. It is possible, then, that in addition to this factor of wide scope, there may be other factors dealing with narrower aspects of job satisfaction, such as pay, working conditions, and so on which do not appear in this analysis. Further research is therefore indicated in order to include sufficient variables to define these hypothesized factors.

Factor II has been called "Knowledge of Employee Benefits," since the greater number of items are concerned with information concerning personnel policies which are related to benefits. The variables which identify this factor are indicated in Table 3.

It seems quite logical that members of the retirement system should be familiar with this information. The negative loading of age on this factor is an artifact, for two reasons. First, a positive weight was given those below 30, and a negative weight those above 30. In addition, all employees over 30 are automatically members of the retirement system, unless physically or otherwise disqualified. It is reasonable to expect women to have more information, since most of the women are clerical workers and the University's Clerical Council furnishes some information in this line that other workers lack. It is reasonable also to expect employees who are married and those of long service to have an interest in, and therefore know more concerning, policies affecting their security. One loading is somewhat anomalous. That is the tendency for workers with smaller amounts of education to know more concerning employee benefits than workers of a higher educational level. The group with the highest rate of turnover is the food service employees, and these are all of low educational level. A possible explanation is that the low education employees are also relatively low in pay, and therefore have fewer personal resources. These workers must then fall back on benefits guaranteed by the University. This has been corroborated by personal interviews.

Summary

The responses of 379 non-supervisory non-academic employees of the University of Illinois to 34 questionnaire items were intercorrelated and factor-analyzed. Of eight factors rotated, only the first two were clearly defined. These were called "Job Satisfaction" and "Knowledge of Employee Benefits."

Received September 27, 1951.

References

1. Carlson, H. B. Attitudes of undergraduate students. *J. soc. Psychol.*, 1934, **5**, 202–213.
2. Chesire, L., Saffir, N., and Thurstone, L. L. *Computing diagrams for the tetrachoric correlation coefficient.* Chicago: Univ. of Chicago Bookstore, 1933.
3. Ferguson, L. W. Primary social attitudes. *J. Psychol.*, 1939, **8**, 217–223.
4. Ferguson, L. W. The stability of the primary social attitudes: I. Religionism and humanitarianism. *J. Psychol.*, 1941, **12**, 283–288.
5. Ford, R. N., and Henderson, D. E. V. A multiple factor analysis of Ford's White-Negro Experience Scales. *Soc. Forces*, 1942, **21**, 28–34.
6. Ford, S. M. *Attitudes of non-academic employees, University of Illinois.* Unpublished Master's Thesis, University of Illinois, 1948.
7. Ghiselli, E. E., and Brown, C. W. *Personnel and industrial psychology.* New York: McGraw-Hill, 1948.
8. Hayes, S. P., Jr. The interrelations of political attitudes: III. General factors in political issues. *J. soc. Psychol.*, 1939, **10**, 379–398.
9. Heist, Arlene, and Heist, P. *An attitude survey of the nonacademic personnel of the University of Illinois, 1948.* Unpublished report on file at the Office of Nonacademic Personnel, University of Illinois, Urbana, Illinois.
10. Hull, R. L., and Kolstad, A. Morale on the job. In Watson, G., *Civilian morale.* New York: Reynal and Hitchcock, 1942.
11. Stagner, R., and Katzoff, E. T. Fascist attitudes: Factor analysis of item correlations. *J. soc. Psychol.*, 1942, **16**, 3–9.
12. Strong, E. K. *Psychological aspects of business.* New York: McGraw-Hill, 1938.
13. Thurstone, L. L. *Multiple factor analysis.* Chicago: Univ. of Chicago Press, 1947.
14. Vetter, G. B. What makes attitudes and opinions "liberal" or "conservative"? *J. abnorm. soc. Psychol.*, 1947, **42**, 125–130.
15. Watson, Goodwin (Ed.). *Civilian morale.* New York: Reynal and Hitchcock, 1942.

A Factor Analysis of Terman and Miles' M-F Test

C. Fenton Ford, Jr.* and Leona E. Tyler

University of Oregon

No area in differential psychology has been explored more extensively than that of sex differences. Large and significant differences between many differently selected samples of males and females have been obtained using tests of mechanical and clerical aptitude, personality questionnaires, and interest inventories. The most thorough single study was that of Terman and Miles.[1] Their *Attitude-Interest Analysis Blank* (M-F Test) is a questionnaire made up entirely of items upon which the sexes differ significantly. The total score for an individual gives a masculinity or femininity rating which can be compared with norms for a wide variety of groups of men and women.

One question needing to be answered with regard to this inventory and others of similar design has to do with the possible multidimensionality of the "masculinity" or "femininity" it measures. Is "masculinity" a unitary trait or a composite of several?

It is this question to which the present study was directed. The subjects were ninth grade students attending junior high schools in Eugene, Oregon, 153 girls and 157 boys. Classes to be tested were selected at the convenience of the teaching staffs of the schools involved, and there is no reason to suppose that they are not representative of the school population of that grade. The Terman and Miles *Attitude-Interest Analysis Blank* was given and scored as directed in the authors' manual. In order to furnish a larger number of part scores and avoid heterogeneity of content within any subtest, scores were grouped into fourteen subscores rather than the seven indicated on the blank. They were as follows: (1) Word Association; (2) Ink Blot Identification; (3) Information; (4) Things That Arouse Anger; (5) Things That Arouse Fear; (6) Things That Arouse Disgust; (7) Things That Arouse Pity; (8) Ethical Attitudes; (9) Interests; (10) Book Likes; (11) Activity Preferences; (12) Famous Persons; (13) Opinions; and (14) Introvertive Response.

Split-half reliability coefficients and percentages of overlapping for the two sex groups were calculated for all the subtests separately. Exercises 2, Ink Blot Identification, and 13, Opinions, were omitted from further computations because of low reliability and poor discriminating power.

Separate correlation matrices were developed for the two sex groups and factor analyzed, using Thurstone's centroid method with orthogonal rotation to simple structure. For the masculinity matrix, four successive repetitions of the analysis were made to stabilize the communalities. For the femininity matrix three repetitions were sufficient.

For the boys, the result of the analysis is given by the rotated factor matrix included in Table 1.

The first factor seems to be an emotional characteristic which might be named "Toughness" or "Insensitivity" since the kind of response scoring masculine on the subtests in question characterizes a person not easily moved to anger, disgust, or pity. The second factor is obviously an interest factor. Just why Exercises 3 and 12 show zero loadings on both factors after rotation cannot be ascertained from the data. They perhaps represent further factors not found to any extent in any other subtests included in the battery. The adjusted communalities for both tests were practically zero.

The rotated factor matrix for the girls is shown in Table 2.

For girls also the first factor would seem to be of an emotional nature and can be labeled "Sensitivity" since the feminine responses to items in the subtests involved are in the direction of arousal of the emotions. Factor II

* Mr. Ford died of a serious illness on October 5, 1950, about two months after completing this study. It has been summarized for publication by Leona E. Tyler. The unpublished M.S. thesis in which the results were reported is available through the library of the University of Oregon.

[1] Terman, L. M., and Miles, Catherine C. *Sex and personality: Studies in masculinity and femininity.* New York: McGraw-Hill, 1936.

Table 1

Centroid and Rotated Factor Loadings, Masculinity, $N = 157$ Boys

Exercise	Centroid		Rotated*	
	Factor I	Factor II	Factor I	Factor II
1. Word Association	.23	−.13	.1	.2
3. Information	−.06	.02	.0	.0
4. Anger	.51	.34	.6	.0
5. Fear	.29	.19	.3	.0
6. Disgust	.74	.36	.8	.1
7. Pity	.58	.31	.6	.0
8. Ethical Attitudes	.70	.22	.7	.2
9. Interests	.50	−.37	.2	.6
10. Books	.39	−.49	.0	.6
11. Activity Preference	.41	−.53	.0	.7
12. Famous People	−.02	−.04	.0	.0
14. Introvertive Response	.37	.08	.3	.1

* Angle of rotation = 32° 07′.

is again primarily an interest factor. Factor III is more difficult to interpret, but in the light of Terman and Miles' description of the nature of characteristically feminine responses to subtests 1, 4, and 9, it appears to be primarily a social role factor reflecting a girl's awareness of the part she is expected to play in our culture. Why Exercise 5, Things That Arouse Fear, should have a negative loading is not clear, unless some concept of bravery, normally thought to be a masculine trait, enters into the stereotype of the ideal woman. In general, Exercise 5 gives somewhat inconsistent results in this study, with a factor loading in the masculinity matrix considerably lower than the other subtests on

Table 2

Centroid and Rotated Factor Loadings, Femininity, $N = 153$ Girls

Exercise	Centroid			Rotated*		
	Factor I	Factor II	Factor III	Factor I	Factor II	Factor III
1. Word Association	.40	−.35	.53	.1	.0	.8
3. Information	.22	.04	.11	.2	.0	.2
4. Anger	.70	.14	.42	.6	−.1	.6
5. Fear	.33	.41	−.45	.5	.3	−.4
6. Disgust	.52	.35	−.08	.6	.1	.0
7. Pity	.39	.40	.07	.6	−.1	.0
8. Ethical Attitudes	.51	.19	.00	.5	.1	.1
9. Interests	.45	−.38	−.06	.1	.4	.4
10. Books	.27	−.34	−.26	.0	.5	.1
11. Activity Preferences	.41	−.53	−.21	.0	.6	.3
12. Famous People	.27	.03	.05	.2	.1	.2
14. Introvertive Response	.24	.05	−.13	.1	.2	.0

* Angle of first rotation (Planes I–II) = 41° 40′; and angle of second rotation (Planes II′–III) = 43° 22′.

emotional characteristics, and an appreciable loading on Factor II as well as Factor I in the femininity matrix.

On the basis of these results we are justified in concluding that psychological masculinity-femininity is not a unitary trait. It has at least two dimensions for both sexes, representing emotional characteristics and interests respectively. There is another possible third dimension for females having to do with acceptance of a feminine social role. Possibly the inclusion of a wider variety of M-F tests in a factor analysis would identify still more factors. As a consequence, it would appear that one should scrutinize closely the *pattern* as well as the extent of M-F deviations represented by scores on any of our currently used tests in this area before making judgments about individual cases. Terman and Miles' Exercise 4, *Emotional and Ethical Attitudes*, and Exercise 5, *Interests*, represent fairly well the two principal factors which this study has identified from the blank as a whole.

Received July 30, 1951.

Vocational Interests of Retired YMCA Secretaries *

Wallace A. Verburg

University of Kansas

Strong's studies (2) of vocational interests were limited primarily to students and employed individuals under 60 years of age. This investigation is concerned with the vocational interests of retired YMCA general secretaries, some of whom were included in Strong's criterion group in 1927.

The general hypothesis of this research was that there would be no significant difference between the vocational interests of retired and employed members of an occupation.

Procedures

The *Strong Vocational Interest Blank* was mailed to 131 retired and 134 active YMCA general secretaries who lived in all parts of the United States. Retirement at 60 is customary in the YMCA, and is compulsory at 65. The return of usable *Blanks* was 66 per cent. No differences obtained between frequency of response from employed and retired secretaries, but in the retired group persons 66 years of age and below tended to respond more frequently than those above 66 years of age. The difference, tested by the chi square technique, was significant at the .02 level. Britton (1) noted no difference in frequency of response to an attitude inventory by retired YMCA secretaries above and below 70 years of age.

Respondents in the retired group in this study had a mean age of 66, 14.7 years of education, an average of 37 years of service in the YMCA, and had been retired for an average of 4.4 years. Respondents in the active group had an average age of 56 years, 15.6 years of education, and 32.3 years of service. Generally, retired YMCA secretaries report themselves to be healthy, financially secure, and sociable. They were satisfied with their jobs before retirement and are satisfied in retirement.

Blanks were machine scored for twelve occupational scales, six of which are "typical" of Strong's groups, and three nonoccupational scales. Pretest (1927) scores for two subgroups, 20 retired (Group RS) and 28 active (Group AS) YMCA secretaries were available from Strong's criterion groups. These data permitted a check on change in interests over a 24-year period for these two groups.

Results

Results of the 1927 and 1951 tests for the two small subgroups (RS and AS) are presented in Table 1.

No difference between mean scores for Groups AS and RS, on each of the fifteen scales in 1927, is significant at or beyond the .05 level. Neither is there any significant difference between the 1951 scores for these two groups.

Two numerical differences between the 1927 and 1951 scores for Group AS on the social science teacher and the public administrator scales are significant beyond the .01 level.[1] Group RS has smaller numerical but also significant changes on these two scales. These two differences are large enough to change the equivalent ratings for both Groups AS and RS from B's to A's. Other differences, although some are statistically significant, do not constitute meaningful changes. For instance, an increase from 13.1 to 16.1 in standard score on the carpenter scale for Group AS, while significant at the .05 level, does not constitute a meaningful change in interest. Both scores are equivalent to C ratings which indicate a lack of interest in the occupation.

It can be noted that generally the active secretaries have greater numerical changes in mean scores than the retired secretaries. However, while the active group gained .8 of a standard score on the accountant scale, the retired group gained 4 points in 24 years.

* The major part of an Ed.D. thesis done by the writer in 1951 (4) under the supervision of Dr. Donald E. Super, Teachers College, Columbia University.

[1] Since 1927 and 1951 scores are correlated, mean changes in standard scores were computed and these changes were tested for significance.

Table 1

Interest Scores of Retired (RS) and Active (AS) YMCA Secretaries Tested in 1927 and Retested in 1951

	Mean Scores				Standard Deviations			
	1927		1951		1927		1951	
Scale / Group: / N:	RS 20	AS 28	RS 20	AS 28	RS 20	AS 28	RS 20	AS 28
Y secy.	51.1	46.9	49.1	48.6	11.7	10.7	11.6	11.1
Y phys.	42.2	39.6	42.2	40.2	12.1	11.5	11.5	8.9
S. sci. tch.	40.7	38.8	46.0	46.1	9.9	9.1	10.5	9.2
Minister	42.2	37.2	42.4	41.2	11.6	11.9	11.5	12.5
Pub. adm.	40.1	38.4	45.2	45.7	10.4	9.4	10.2	9.3
Life ins.	37.9	36.0	37.1	34.9	9.8	10.2	8.3	11.1
Acct.	28.9	31.1	32.9	31.9	11.5	12.5	9.8	9.5
Lawyer	29.6	30.7	30.2	31.6	9.6	9.2	10.0	8.3
Farmer	29.3	28.0	29.6	30.2	10.1	9.4	9.0	8.7
Physician	22.5	21.4	18.7	19.6	6.1	10.8	6.1	7.9
Chemist	14.2	16.2	13.0	14.8	10.4	12.3	10.0	13.2
Carpenter	16.6	13.1	18.5	16.1	12.2	11.0	12.8	12.1
OL	59.0	59.7	57.2	57.3	8.0	7.3	7.6	7.0
MF	41.5	40.9	41.2	41.2	9.3	8.7	8.1	7.2
IM			60.9	59.4			6.7	4.5

Gains for both groups are in the opposite direction from that anticipated (**2**, pp. 275–276). No change, however, for the retired group can be explained as an effect of retirement. This conclusion is supported further when 1951 scores of larger groups of retired (Group R) and active (Group A) secretaries are compared in Table 2.

The similarity of mean scores for these groups is striking. Also, mean scores for these persons, who were not members of Strong's criterion group, are very similar to those in Table 1. There are no significant differences between 1951 scores on any scale for these two groups and their scores are not significantly different from comparable scores of the smaller groups in Table 1. When standard scores of the two retired groups (R and RS) are combined and scores of the two active groups (A and AS) are also pooled, the greatest difference between combined active and retired groups on the minister scale is significant at the .05 level only. Retired persons scored 3.6 higher on the minister scale which may be partially accounted for by the fact that individuals with interests similar to those of ministers tend to score higher on the minister scale with increased age (**2**, pp. 274–276). Also, the retired secretaries are older and served in a more evangelical Y. This difference cannot be

Table 2

Interest Scores of Retired (Group R) and Active (Group A) YMCA Secretaries Tested in 1951

	Mean Scores		Standard Deviations	
Scale / Group: / N:	R 61	A 65	R 61	A 65
Y secretary	48.3	47.0	10.0	10.4
Y physical director	41.5	41.8	10.3	11.8
Social science teacher	44.3	44.8	10.4	10.4
Minister	42.0	37.3	12.5	10.7
Public administrator	41.3	43.7	9.3	10.0
Life insurance salesman	38.0	38.8	9.5	9.6
Accountant	32.5	29.7	9.7	10.7
Lawyer	29.6	29.6	8.5	7.7
Farmer	29.0	29.3	9.7	9.1
Physician	21.5	20.3	9.5	9.4
Chemist	13.3	13.2	11.0	11.3
Carpenter	17.5	16.2	13.5	12.6
OL	56.6	57.4	7.9	7.7
MF	39.3	41.7	8.3	8.1
IM	59.7	58.9	6.2	6.0

interpreted as being very meaningful since Group AS has a rating which is equivalent to a high B and Group RS has a low B+ rating.

Retired YMCA secretaries were classified according to the similarity to YMCA work of jobs they held in retirement. The group ($N = 28$) that was employed in jobs dissimilar to their previous occupation could not be differentiated on the YMCA secretary scale from the group ($N = 38$) that continued in jobs after retirement which were similar to YMCA work. It was possible to score only eleven *Blanks* for dissimilar occupations since scales for all of the 28 dissimilar occupations were not available. These eleven individuals had an average mean score equivalent to a B+ rating on their dissimilar occupations. Their mean YMCA secretary scale score was equivalent to an A rating. Apparently, vocational interest scores for YMCA secretaries are maintained even though individuals may pursue an occupation in retirement which is rated as being dissimilar to their former work.

The supposition that retired persons would score lower than active secretaries on the Occupational Level, Interest Maturity and Masculinity-Femininity scales was not borne out by this study. Strong (**2**) and Terman and Miles (**3**) have reported that men tend to become more feminine in interests with advanced age. There were no significant differences between M-F scores for the combined groups of retired and active YMCA secretaries in this study but when the retired and active group M-F scores were compared with those of groups of younger men (**2**, p. 231), they were significantly lower (more feminine). Mean scores for the active YMCA secretaries were also significantly lower (.05 level) and retired secretaries lower (.01 level) than for a group of 317 men between the ages of 50 to 59 from Strong's mixed criterion groups. However, 61 individuals from Strong's men-in-general group between the ages of 50 to 59 scored significantly lower (.0001 level) than both the active and retired groups of YMCA secretaries in this study.

Summary

Vocational interests of 131 retired and 134 active YMCA general secretaries were measured by the *Strong Vocational Interest Blank*. No significant differences obtained between 1951 mean standard scores on 15 interest scales for retired and active groups. A tendency for both employed and retired YMCA secretaries to score slightly higher over a 24-year period on the social science teacher and public administrator scale was noted. Retired secretaries had changes in scores on most scales which were numerically less than those for employed secretaries but their 1951 scores were very similar to those of employed secretaries. The observation that M-F scores for groups of older men, both active and retired, tend to be more feminine than those for younger men supports previous findings.

1. Vocational interest scores, as indicated by 12 occupational scales, for a group of YMCA general secretaries are maintained in reitrement at least beyond the age of 70 years.

2. Vocational interest scores for groups of YMCA general secretaries, both retired and active, remain relatively constant over a 24-year period.

3. Occupational Level, Interest Maturity, and Masculinity-Femininity scores apparently are not affected by retirement. M-F scores for groups of retired and active secretaries are lower than M-F scores of younger men; however, YMCA secretaries in later maturity have more masculine interests than men-in-general of comparable age.

4. Younger retired YMCA secretaries tend to return a questionnaire similar to the *Strong Blank* more frequently than older retired secretaries.

Received August 1, 1951.

References

1. Britton, Jean O. *A study of the adjustment of retired YMCA secretaries.* Unpublished Ph.D. dissertation, Univ. of Chicago, 1949.
2. Strong, E. K., Jr. *Vocational interests of men and women.* Stanford: Stanford Univ. Press, 1943.
3. Terman, L. M., and Miles, Catherine C. *Sex and personality.* New York: McGraw-Hill, 1936.
4. Verburg, W. A. *A study of the effects of retirement on the interests of YMCA secretaries as indicated by the Vocational Interest Blank.* Unpublished Ed.D. project, Teachers College, Columbia Univ., 1951.

Predicting Success in Law School *

Ralph F. Berdie and Wilbur L. Layton

Student Counseling Bureau, University of Minnesota

Professional schools are among the most frequent users of prediction indices for purposes of selection. When facilities for training are limited and laboratories and equipment expensive, only those students who show the greatest probability of succeeding are wanted.

The prevailing policies in many law schools, however, allow any student to enter who has completed the minimum number of years, two, three, or four, of pre-legal work. Selection then takes place, not prior to admission, but during the first years of legal training, and those who fail have the compensation of knowing they had a chance. But those who fail also run the risk of a serious traumatic experience to say nothing of the financial wastes involved.

Predictive tests have proved to be of much value to counselors and instructors working with law students, both those who fail and those who do not. A student whose scores indicate he has ability to do well, but who receives poor grades during his first year, can be encouraged and assisted to develop proper study methods or to discover other reasons underlying his poor work. A failing student with low test scores can be encouraged to re-evaluate his capacities for this training and to consider alternative training programs and occupations. In this situation, scholastic prediction does not cease when a student is admitted. It must continue as long as the student is to have counseling and personalized instruction.

Purpose

The purpose of this study was to determine the effectiveness of predictive data available for freshmen in law school at the end of their first quarter in residence. The predictive data were based on tests, pre-legal grades, and a score on a "practice examination" given by the law school at the end of the first quarter. The criterion was the over-all grade of the student at the end of his freshman year in law school, based on grades in his four freshman-year courses.

Method

The 207 freshmen entering the law school of the University of Minnesota in the fall of 1947 were given a battery of tests during the first week of the quarter. These tests included the Miller Analogies Test, Form G, and the Iowa Legal Aptitude Test, consisting of seven parts—analogies, reasoning, opposites, relevancy, mixed relations, memory, and legal information. A total score was also determined. Other data available for these freshmen included scores on the ACE Psychol. Exam., 1937 form, taken for the most part during the senior year in high school; high school percentile rank; pre-legal honor point ratio; and grades based on the first quarter's work. These latter grades are based on comprehensive and lengthy essay type examinations. Students are not dropped from school on the basis of grades on these tests since these first quarter examinations are regarded as exploratory practice. Grades on these first quarter examinations had a correlation of .56 with first year grades, based on the comprehensive examinations covering the work of the entire freshman year.[1]

At the end of the first quarter, correlations were computed between test scores and grades for that quarter. At the end of the year, zero order and multiple correlations were computed between test scores, first quarter grades, and first year grades.

The multiple regression equations determined on the group of 1947 freshmen were used to predict first year grades for the fresh-

* Appreciation is expressed to Dean Maynard E. Pirsig and Professors Robert C. McClure and Stanley V. Kinyon of the School of Law at the University of Minnesota for their interest in and cooperation with this project.

[1] Although the predicted first year grades are available to the students and faculty at the close of the first quarter, the final examinations are graded anonymously at the end of the freshman year and the authors are convinced the criterion is not contaminated.

men who entered in the falls of 1948 and 1949. These predictions were also made at the end of the fall quarter in each case. The predicted grades were compared with the obtained grades to determine the "shrinkage."

Results

The zero order correlations between the predictive data and first quarter grades and first year grades are presented in Table 1.

The regression equation for predicting first quarter grades is:

$$\hat{Y}_1 = 6.3510X_2 + .1195X_3 + 52.2233$$

and the multiple $R = .49$,

where

$\hat{Y}_1$ = Fall quarter grades
X_2 = Pre-law honor point ratio
X_3 = Legal information

The regression equation for predicting first year grades is:

$$\hat{Y}_2 = .3166X_1 + .0542X_2 + 3.5544X_3 - .3848X_4 + .4738X_5 + .0399X_6 - .2183X_7 + 38.1640,$$

Table 1

Correlations Between Test Scores and Grades at the End of the First Quarter and at the End of the First Year for Law Freshmen Entering in 1947

Predictor	Correlation with 1st Quarter Grades	Correlation with 1st Year Grades
	$N = 99$	$N = 99$*
Miller Analogies Test, Form G	.22	.40
ACE Psychol. Exam.	.23	.49
High school percentile rank	.23	.39
Pre-law honor point ratio	.48	.55
Iowa Legal Aptitude Test—		
Total	.33	.47
Analogies	.32	.49
Reasoning	.17	.38
Opposites	.24	.33
Relevancy	.21	.33
Mixed relations	.26	.30
Memory	.14	.16
Legal information	.27	.27
First quarter grades	—	.56

* Complete data were available on 99 of the 207 freshmen.

where

$\hat{Y}_2$ = First year grades
X_1 = First quarter grades
X_2 = A.C.E. (percentile rank)
X_3 = Pre-law honor point ratio
X_4 = Mixed relations
X_5 = Analogies (Iowa)
X_6 = High school rank
X_7 = Memory

The multiple R, using these seven variables to predict freshman grades, was .75. Combining the variables to predict first year grades increased the correlation as the multiple R of .75 is higher than the highest zero order correlation of .56 between first year grades and first quarter grades. In predicting first quarter grades, however, combining the variables, which resulted in a multiple R of .49 did not provide significantly better prediction than was obtained by using pre-law honor point ratio alone, where the zero order correlation was .48.

Using Hull's method of expanding weights to allow the range of predicted grades to approach the range of obtained grades, the equation for predicting first quarter grades becomes:

$$\hat{Y}_1 = 12.95X_2 + .2438X_3 + 44.18$$

and the equation for predicting first year grades becomes:

$$\hat{Y}_2 = .4211X_1 + .0721X_2 + 4.7274X_3 - .5118X_4 + .6302X_5 + .0531X_6 - .2903X_7 + 26.85.$$

For several students, high school percentile ranks, pre-legal grades and scores on the ACE Psychol. Exam. were not available so a prediction equation based on four predictors was developed.

$$\hat{Y}_2 - .6246X_1 - .4197X_4 + .7385X_5 + .3116X_8 + 13.44,$$

where X_8 = reasoning (Iowa test). The multiple R provided by this combination was .68.

In the falls of 1948 and 1949 the two freshman classes entering those years were given the tests found useful in predicting grades for the 1947 class and at the close of the fall

Table 2

Correlations Between Obtained First Year Grades and Grades Predicted on the Basis of Four and Seven Variable Regression Equations for the Freshmen, 1947–1949

Class Entering Year of:	*R* Based on 4 Variable Prediction	*R* Based on 7 Variable Prediction
1947	.68	.75
1948	.67	.59
1949	.79	.73

quarters, first year grades were predicted, using both the seven and the four variable prediction equations presented earlier, those with the expanded weights. The obtained first year grades were then correlated with the predicted grades. These correlations are shown in Table 2.

For the 1948 class, the zero-order correlation between obtained grades and predicted grades, using the seven variable equation, was .59, and using the four variable equation, .67. Similar correlations for the 1949 group were .73 and .79. These are to be compared to the originally obtained multiple *R*'s of .75 and .68. The means and standard deviations for these three years are found in Table 3.

Table 4 presents the distribution of predicted grades as compared to the obtained grades for each of the three years. Of the freshmen entering in 1947, 1 per cent had predicted grades of 60 to 69 and actually obtained grades of 80–89, whereas of this total class, 5 per cent had predicted grades of 80–89 and actually

Table 4

Per Cent of Students Each Year with Predicted and Obtained Grades, Predictions Being Based on Four Variable Equation

Obtained Grades	Year	Predicted Grades 50–59	60–69	70–79	80–89	Total
	1949			7	11	18
80–89	1948		1	7	6	14
	1947		1	13	5	19
	1949	1	11	30	6	48
70–79	1948	3	22	28	3	56
	1947	1	13	37	1	52
	1949	4	17	6		27
60–69	1948	7	13	6		26
	1947	4	19	4	1	28
	1949	2	4	1		7
50–59	1948	2	3			5
	1947	2	1			3
Totals	1949	7	32	44	17	
	1948	12	39	41	9	
	1947	7	34	54	7	

Table 3

Means and Standard Deviations of Test Scores and Grades for the Law Freshman Classes of 1947, 1948, and 1949

Variable	1947 Group A Mean	SD	*N**	1947 Group B Mean	SD	*N*	1948 Mean	SD	*N*	1949 Mean	SD	*N*
First year grades	72.6	7.0	99	72.4	7.1	163	73.2	7.3	250	71.4	13.7	109
First quarter grades	69.1	7.4	99	69.3	6.2	122	65.0	9.5	286	67.6	7.5	127
A.C.E. (percentile rank)	61.5	26.8	99	60.2	26.6	151	60.2	25.7	211	56.2	25.6	123
Pre-law honor point ratio	1.7	.5	99	1.7	.5	124	1.6	.5	268	1.7	.5	138
Mixed relations (Iowa)	15.8	7.1	99	15.6	7.2	207	14.7	7.3	255	14.5	7.0	145
Analogies (Iowa)	21.5	6.4	99	21.2	6.7	207	20.3	6.7	254	20.5	6.3	145
Memory (Iowa)	18.1	3.2	99	17.8	3.1	207	17.1	3.3	255	18.3	3.4	145
Reasoning (Iowa)	21.6	6.4	99	20.6	6.6	207	20.2	6.3	255	21.1	6.5	144
High school rank	75.3	23.2	99	69.0	24.2	146	65.8	23.8	203	67.6	22.9	123

* For the 1947 class, group A is the group upon which the original regression equations were based and group B consists of the total group for whom a particular measure was available. The *N*'s for the 1948 and 1949 groups vary for the same reason. The differences between these groups are so slight that they probably do not account for the differences in the correlations between predicted and obtained grades for the three years.

Table 5

Percentage of Students Each Year with Various Predicted Grades Who Failed

Predicted Grade	Year		
	1947	1948	1949
Percentage with predicted grade below 60 who actually failed	90	76	87
Percentage with predicted grade of 60 or above who actually failed	26	30	28
Percentage with predicted grade of 70 or above who actually failed	7	12	12
Percentage with predicted grade of 80 or above who actually failed	0	5	0

obtained those grades. For that year, 26 per cent had predicted grades of below 70 and actually obtained grades of below 70, while 15 per cent of that class had predicted grades of below 70 but actually did better than that. Table 5 summarizes some of the information presented in Table 4.

Thus, of 46 students predicted to have grades of 80 or more, only one failed, when the lowest passing grade is 70. Of 43 students predicted to have grades of less than 60, eight passed.

Summary

The predictions studied here differ from those previously reported[2] insofar as a progress measure of success at the end of the first third of the freshman year is used as one of the predictors of academic standing at the close of that year. Pre-legal honor point ratio was found to be the best single predictor of freshman grades here (with the exception of trial examination grades), as has been reported consistently in other studies. High school percentile rank, the ACE Psychol. Exam., and the Iowa Legal Aptitude Test also contributed to effective prediction.

The multiple correlations reported here are not significantly different from others found at Minnesota and at other places attempting to predict grades in law school. The absence of any significant "shrinkage" in predictive efficiency over the three year period is encouraging for those attempting such predictions.

Received August 10, 1951.

[2] Stuitt, D. B., *et al. Predicting success in professional schools.* Washington, D. C.: American Council on Education, 1949.

The ACE Psychological Examination and High School Standing as Predictors of College Success

Norman Frederiksen

Princeton University and Educational Testing Service

and

W. B. Schrader

Educational Testing Service

A study of the college adjustment of veteran and nonveteran students has recently been completed by the Educational Testing Service, with the support of the Carnegie Corporation of New York. The design of this study involved the calculation of validity coefficients for high school standing and the ACE Psychological Examination (the ACPE) for freshman students in a number of colleges and universities. The present article summarizes the findings of these validity studies.

The Groups Studied

The colleges and universities which participated in the study were quite varied. They were so chosen as to include private colleges, state universities, and municipal institutions; coeducational and men's colleges; large universities and relatively small colleges; institutions with great financial resources and less wealthy colleges; and colleges located in large cities and those located in small towns. Geographically they were distributed throughout the country, but with some concentration in the east and midwest. Sixteen institutions were included in the entire investigation, and results for twelve of these colleges will be reported here.

Some care was exercised to ensure that the group of students represented in any particular validity study were reasonably homogeneous. Specifically, the members of such a group were all male students, all attended the same institution, all were enrolled in the same college or division of the institution (e.g., liberal arts or engineering), and all had entered the institution in the fall of 1946. Within each group, veteran and nonveteran students were treated separately. Since, in three of the twelve colleges considered, separate studies were made of students enrolled in different divisions, a total of fifteen groups, each containing both veteran and nonveteran students, were investigated. In each group the validity of the ACE *Psychological Examination* or of some measure of high school standing, or of both, was studied. In each case, the criterion was the first-year college average grade.

Results

The validity coefficients are presented in Table 1; the corresponding means and standard deviations are shown in Table 2. (In general, the means and standard deviations are comparable only for veterans and nonveterans at the same institution, since the scales are not necessarily the same at different institutions. Raw scores on the ACPE were based on various forms of the test, and in some instances percentile scores were converted to standard scores with a mean of 13 and a standard deviation of 4. High school standing was reported on a variety of scales.)

The validity coefficients for the ACPE showed considerable fluctuation from one subgroup to another; much of this variability is presumably ascribable to sampling error, particularly since some of the subgroups are rather small.[1] The median correlation for the twenty-two subgroups of male students where ACPE total score was used is .47. The Thurstones[2] observed, in 1932, that the correlations

[1] For the eleven veteran subgroups, the validity coefficients range from .35 to .60, but it was found that the variation among the coefficients was no greater than would be expected to arise by chance in groups of this size. For the eleven nonveteran subgroups, the range is from .28 to .61; in this case, there are less than 2 chances in 100 that as great or greater variability would arise by chance.

[2] Thurstone, L. L., and Thurstone, Thelma G. The 1931 Psychological Examination. *Educ. Rec.*, 1932, **13**, 121–136.

Table 1

Correlations of ACPE Scores and Measures of High School Standing with First-Year College Average Grade

College	Division	Description of College	Validity Coefficients: ACPE Total Score MV*	ACPE Total Score MN**	H. S. Standing MV*	H. S. Standing MN**	Measure of H. S. Standing Used	*N* MV*	*N* MN**
A	Arts	Men, private, east	—	—	.55	.60	Adjusted Rank	694	531
B	Arts	Men, private, east	—	—	.53	.62	Adjusted Rank	536	570
C	Arts	Coed., private, east	.41	.39	.43	.58	Rank	111	112
D	Arts	Coed., private, south	.35	.55	.52	.65	Average Grade	77	119
E	Arts	Coed., private, midwest	—	—	.49	.65	Rank	164	175
F	Arts	Coed., private, midwest	.51	.45	—	—		283	94
G	Arts	Coed., public, midwest	—	—	.61	.61	Average Grade	466	166
H	Arts	Coed., public, midwest	.49	.34	—	—		111	115
I	Arts	Coed., public, midwest	.57	.29	—	—		83	72
J	Arts	Coed., public, west	.46	.51	.53	.59	Average Grade	433	222
K	Engr.	Coed., public, midwest	.60	.61	.62	.53	Average Grade	271	128
L	Engr.	Coed., public, midwest	.42	.28	.51	.53	Rank	352	98
I	Engr.	Coed., public, midwest	.45	.36	—	—		167	171
K	Agric.	Coed., public, midwest	.52	.48	.56	.68	Average Grade	140	102
H	Bus.	Coed., public, midwest	.55	.57	—	—		142	65
		Median Validity Coefficient	.49	.45	.53	.60			

* Male veterans.
** Male nonveterans.

between ACPE scores and college grades averaged around .50, which is in good agreement with the present findings.

There is a slight tendency for the test to yield higher correlations with grades of veterans than with grades of nonveterans; the median validity coefficients are .49 and .45 respectively.[3] The slight tendency for validity coefficients to be higher for veterans is consistent with the hypothesis that veterans are more uniformly motivated than nonveterans and thus tend more than nonveterans to achieve the grades they are capable of earning.

The correlations of high school standing with first-year grades varied from .43 to .68 in the twenty subgroups where it was used as a predictor; the median validity coefficient is .57.[4] High school record thus furnishes a somewhat more accurate prediction, for these groups, than the ACPE. While the ACPE tended slightly to be more valid for veterans than for nonveterans, the opposite tendency was found for high school standing; the median validity coefficients were about .53 and .60 for veteran and nonveteran males respectively. Out of ten comparisons, the validity of high school standing was greater for nonveterans in eight instances, with one tie. This difference seems reasonable in view of the greater amount of time elapsing between high school and college for the veteran students. High school grades presumably reflect motivational and other nonintellectual factors as well as ability to do academic work; to the extent

[3] When validity coefficients for veterans and nonveterans are compared in the eleven groups, the coefficient is higher for the veterans in seven of the eleven comparisons. In only one instance, however, is the difference significant; the coefficient for veterans is significantly higher at the 5 per cent level for the group of engineers at University L.

[4] For the ten subgroups of veterans and for the ten subgroups of nonveterans, separately, it was found that the variability in the correlation coefficients did not exceed the amount to be expected by chance. When validity coefficients for veterans were compared with those for nonveterans in each of the ten groups separately, the coefficients for the nonveterans were found to be significantly higher (5 per cent level) for students in the liberal arts division in two universities (B and E).

Table 2

Means and Standard Deviations of Predictor and Criterion Variables

College	Division	Subgroup	ACPE		H. S. Standing		First-Year Avg. College Grade	
			Mean	SD	Mean	SD	Mean	SD
A	Arts	MV*	—	—	72.2	7.7	75.1	6.1
		MN**	—	—	76.2	7.0	75.8	6.4
B	Arts	MV	—	—	4.12	.76	4.76	.77
		MN	—	—	4.70	.78	5.07	.83
C	Arts	MV	111	20	14.8	3.3	1.49	.56
		MN	112	20	15.8	3.4	1.50	.64
D	Arts	MV	106	19	3.26	2.04	1.14	.64
		MN	117	20	4.07	2.30	1.40	.75
E	Arts	MV	—	—	15.0	3.0	1.20	.72
		MN	—	—	17.2	3.4	1.43	.71
F	Arts	MV	12.0	3.3	—	—	2.10	.61
		MN	13.1	3.1	—	—	2.05	.60
G	Arts	MV	—	—	2.48	.62	2.20	.68
		MN	—	—	2.74	.64	2.19	.66
H	Arts	MV	111	23	—	—	2.29	.68
		MN	115	20	—	—	2.35	.69
I	Arts	MV	120	23	—	—	3.40	.77
		MN	120	23	—	—	3.18	.80
J	Arts	MV	114	22	2.69	.50	2.39	.62
		MN	114	21	2.93	.51	2.43	.56
K	Engr.	MV	13.7	3.5	2.75	.62	2.39	.63
		MN	14.4	3.5	2.96	.56	2.24	.63
L	Engr.	MV	88.1	21.5	15.6	3.1	1.40	.66
		MN	93.1	22.3	16.4	3.3	1.32	.60
I	Engr.	MV	118	19	—	—	3.34	.74
		MN	115	20	—	—	3.32	.77
K	Agric.	MV	11.9	3.9	2.52	.64	2.30	.60
		MN	11.5	3.5	2.75	.59	2.15	.51
H	Bus.	MV	106	23	—	—	2.34	.65
		MN	107	21	—	—	2.23	.61

* Male veterans.
** Male nonveterans.

that these characteristics have changed more for some veterans than for others, the correlation would be lowered. To the extent that high school grades reflect knowledge and skills directly useful in college work, their predictive effectiveness would be lowered also, since the educational effects of war service were not uniform for all veterans.

The universities which supplied information on high school standing used various measures of relative performance. The three types of measures employed have been designated as

average grade, rank in class, and adjusted rank in class. Of these, average grade suffers from the disadvantage that the various secondary schools use marking systems which differ markedly in form. Rank in class overcomes this difficulty, and is presumably preferable to average grade; but the secondary schools may differ greatly in the calibre of their students, so that being, for example, 75th in a class of 100 does not mean the same thing in all schools. The adjusted rank in class provides a system of corrections to overcome this difficulty, based on past experience with the graduates of the various schools.

The type of measure of high school standing employed at each institution is indicated in Table 1. The correlations shown do not indicate any clear superiority for any particular type. It must be pointed out, however, that the colleges vary in their use of high school standing in selection of students. Some of the colleges admit all graduates of approved high schools, while others do not. Students with poor high school records were not admitted in any of the five groups where rank or adjusted rank was used. The curtailment in range produced by use of the measure in selection obviously will affect the size of the correlation obtained; the relative merits of the different measures of high school standing therefore cannot be appraised from these findings. It may be concluded only that all the colleges utilizing high school record are employing it in a form which is quite effective as a predictor of freshman average grade for its own students.

Table 3

Multiple Correlations of ACPE Total Score and High School Standing with First-Year College Average Grade

College	Division	Subgroup	*R*
C	(Arts)	Male veterans	.53
		Male nonveterans	.62
D	(Arts)	Male veterans	.59
		Male nonveterans	.74
J	(Arts)	Male veterans	.60
		Male nonveterans	.65
K	(Engr.)	Male veterans	.72
		Male nonveterans	.76
L	(Engr.)	Male veterans	.56
		Male nonveterans	.53
K	(Agric.)	Male veterans	.65
		Male nonveterans	.70
Median		Male veterans	.60
		Male nonveterans	.68

In six of the groups, validity data are available for both the ACPE and high school standing, thus affording twelve direct comparisons of the two predictors. For veterans in these six groups, the median validity of the test is .44 and of high school standing about .52; the comparable medians for nonveterans are about .50 and .58. In eleven of the twelve comparisons, the validity coefficient was higher for high school standing; such consistency in the direction of the difference would be expected to occur by chance fewer than once in a hundred times.[5] The finding of a higher validity coefficient for high school standing than for scores on a single aptitude test agrees with the typical findings of previous comparisons of this kind.

It is, however, more proper to think of the test and high school record as joint members of a predictive team than as competitors. The effectiveness of the team is shown in Table 3, which presents the multiple correlations of ACPE total score and high school record with first-year average grade for the six institutions employing both measures. For veterans, the *R*'s range from .53 to .72 with a median of about .60; for nonveterans the range is .53 to .76 with a median at about .68. The use of the two predictors in combination thus furnishes a better basis for prediction than either measure used alone. The magnitude of the correlation is great enough to indicate that the combined measures provide a really useful prediction of how well a particular student is likely to succeed in his freshman year.

[5] When the differences between validity coefficients were tested for statistical significance, high school standing had a significantly higher validity in four subgroups. One difference was significant at the 1 per cent level (nonveteran engineers at University L); and three were significant at the 5 per cent level (veteran engineers at L, nonveterans in arts at C and in agriculture at K).

Summary

1. The median correlation of the ACE *Psychological Examination* with first-year college grades was .47; the median correlation of high school standing with the same criterion was .57.

2. The validity of the ACPE tends to be slightly greater for male veterans than for male nonveterans, while the validity of high school standing tends to be greater for nonveterans. The median validity coefficients for the ACPE, based on veteran and nonveteran students, were .49 and .45 respectively. The comparable median validity coefficients for high school standing were .53 and .60.

3. The use of a weighted composite of ACPE total score and high school standing provides a useful prediction of freshman average grade; the median multiple correlation coefficient for veteran students was found to be .60 and for nonveterans .68.

Received August 6, 1951.

Education and Prediction of Military School Success [1]

Virginia Zachert and Abraham S. Levine

USAF Training Command, Human Resources Research Center, Personnel Research Laboratory, Lackland Air Force Base, San Antonio, Texas

An attempt was made to explore the possibilities of using years of formal education as an additional predictive variable in various Air Force training programs. The procedure involved comparing the validities of various Air Force classification test composites with the validities resulting from using amount of education as a predictor. For the purposes of this study, none of the obtained validity coefficients were corrected for restriction of range due to initial screening on the basis of either aptitude test scores or education. In most cases, the samples were screened on the basis of aptitude test variables, but since years of education and aptitude test performance are substantially correlated, it must be inferred that the populations sampled in this study were more homogeneous than the general population in both the aptitude and educational dimensions. For the most part, the enlisted airmen in this study were recruited from the upper 60% of the population, while the aircrew trainees were more representative of the college student population.

All of the tests utilized in this investigation are included in the Airman Classification Battery or Aircrew Classification Battery, which were developed by aviation psychologists during or following World War II. The educational attainment data were obtained from forms completed by the basic airmen. In each comparison the total number of available cases was used.

World War II Validities

A survey of aircrew studies during World War II (**2**) showed that the validity for educational level was always below that for the appropriate stanine.[2] For an experimental group of 1,300 men admitted to pilot training without any requirements as to aptitude or personality, educational level had a validity of .20 as compared to the validity of .64 for the Pilot Stanine. The criterion measure employed was graduation vs. elimination in Primary Flying School. However, for a group initially screened with the Army Air Forces Qualifying Examination, the validity of the Pilot Stanine was .55 and education dropped to .02. Apparently, virtually all of the valid variance in the education predictor was accounted for by the screening test. For a sample of nearly 2,000 advanced trainees, selected on the Navigator Stanine, the validity for education was .11 while the Navigator Stanine validity was between .50 and .60. For bombardier trainees, the validity of the Bombardier Stanine ranged from .10 to .30, and the education validity was about .05.

Postwar Validities (Aircrew)

Since October 1947, applicants for pilot training were required to present a minimum of two years of college education in addition to passing a screening examination. Thus, with respect to educational level the postwar pilot trainees represented a somewhat more restricted group than the World War II cadets. For all 1949 basic pilot classes (total $N = 1596$), the biserial validity of the Pilot Stanine is .54, and the educational level validity is .04 (**1**). These validities are consistent with those obtained in World War II studies on screened groups.

Postwar Studies (Airmen)

Table 1 presents validities obtained for the first form of the Airman Classification Battery and comparative validities obtained for years of education (**3**). The table also shows the respective contributions of education

[1] The views expressed in this article are those of the authors and do not necessarily represent the official views of the United States Air Force.

[2] By appropriate stanine is meant the one designed to measure success in the particular type of training referred to; thus, the appropriate stanine for pilot trainees is the Pilot Stanine. A stanine is a composite score expressed in nine-point standard score units with a mean of 5 and a standard deviation of 2. The lowest possible stanine is 1 and the highest is 9.

Table 1

Comparison of Airman Classification Battery (AC–1A) and Education for Prediction of Final Technical School Grades

School	*N*	Multiple *R** of Tests in ACB	Education Validity	Ed. Adds to *R*
Clerk-Typist	1199–4715	.68	.58	.05
Stenographer	57–197	.54	.42	.02
Aircraft Welder	43–145	.77	.28	.00
Sheet Metal Worker	168–496	.62	.31	.00
Carpenter	222–881	.49	.29	.01
Electrician	116–412	.61	.36	.00
Fabric and Dope Mechanic	49–107	.44	.35	.02
Machinist	54–232	.63	.38	.01
Parachute Rigger	35–155	.57	.35	.03
Plumber	48–228	.49	.22	.00
Sheet Metal Worker	41–161	.60	.44	.04
Radar Mechanic	52–275	.76	.39	.00
Radio Mechanic	74–195	.75	.61	.06
Remote Control Turret Mechanic	108–304	.56	.30	.00
Construction Equipment Operator	174–537	.42	.12	.00
Firefighter and Crash Rescueman	45–199	.65	.26	.00
Airplane Electrical Mechanic	45–580	.59	.25	.00
A & E Mechanic	243–269	.62	—**	.06
A & E Mechanic (Conventional)	722–1994	.79	.50	.01
A & E Mechanic (Jet)	538–1641	.64	.38	.00
Airplane Propeller Mechanic	53–123	.77	.40	.00
Airplane Instrument Mechanic	87–168	.48	.17	.00
Auto Mechanic Technician	88–413	.58	.22	.00
Diesel Mechanic	44–235	.33	.06	—
Engineman, Operating	86–319	.54	−.04	—
Primary Armament Technician	255–282	.60	.25	.02
Control Tower Operator	68–740	.58	.31	.00
Radio Operator AACS	92–101	.45	—	—
Radio Operator (High Speed Manual)	141–243	.58	.36	.00
Draftsman	121–433	.54	.30	.00
Photographer	52–311	.61	.35	.01
Weather Observer	190–559	.65	.45	.01

* These multiple validity coefficients are based on the best combination of tests in the battery. When the tests are cross-validated, the *R*'s reported above should show some shrinkage.

** For *N* of 23, Educ. Valid. equals .64.

to the multiple *R* of the battery for predicting final grades in various Air Force technical schools. Validation data on these 32 technical schools indicate that the ability tests measure nearly all the valid variance to be found in the educational level variable as well as much variance not accounted for by years of education. The average increase in the multiple correlation when education is added is .012, and the most it adds is .06. Although in many cases the validity for educational level is above .50, the multiple *R* for the tests is at least .15 higher.

Summary

1. Years of education apparently adds little to the predictive efficiency of the Aircrew Classification Battery for various categories of aircrew training. Neither does it appreciably improve multiple prediction of most airmen technical school criteria.

2. Educational level does appear to contribute effectively to the prediction of a few airmen technical school criteria, e. g., Clerk-Typist, Radio Operator. Cross-validity data should be obtained on these schools in order to verify these findings. If they are confirmed, it might prove feasible to weight amount of education in terms of its contribution to a valid battery of tests.

Received August 8, 1951.

References

1. Dailey, J. T., and Gragg, D. B. *Postwar research on the classification of aircrew.* Research Bulletin 49–2, Directorate of Personnel Research, Human Resources Research Center, Lackland Air Force Base, November 1949.
2. Dubois, P. H. *The classification program.* Army Air Forces Aviation Psychology Program Research Reports, No. 2. Washington: 1947.
3. Zachert, Virginia. *Comparison of aptitude tests and years of education as predictors of military success.* Research Note 51–9, Directorate of Personnel Research, Human Resources Research Center, Lackland Air Force Base, April 1951.

The Use of Levers in Making Settings on a Linear Scale *

William Leroy Jenkins and Merritt W. Olson

Lehigh University

Various factors in the use of knobs for making settings on a linear scale have been investigated in three previous studies (**1**, **2**, **3**). The most significant factor turns out to be the ratio between the movement of the pointer along the scale and the rotation of the knob. For relatively fine tolerances (.007 in. and thereabouts), the optimal ratio is one or two inches of pointer movement for one complete turn of the knob.

With much coarser tolerances, it would seem possible that a lever might permit more rapid settings than a knob. The present study is concerned with a comparison between levers and knobs and an attempt to define the significant factors in the use of levers.

Apparatus and Procedure

The apparatus and procedures were essentially the same as those described for the previous studies with knobs (**1**, **2**, **3**). The subject *matches* the position of a lighted insert in a black bakelite scale with a pointer controlled by a knob or lever turning in a plane parallel to the movement of the pointer. Some minor modifications were made in the nature of the scales and in the controlling mechanism for the various experiments.

The permitted error tolerance is determined by the width of the pointer in relation to the width of the lighted insert. Two tolerances were employed: .016 and .100 in.

Time was measured separately for travel to the approximate location and for making the final adjustment. For simplicity, however, results are reported only in terms of the total time.

Results[1]

Experiment 1. The apparatus as used previously in the knob studies was employed with six and twelve-inch levers in comparison with a $2\frac{3}{4}$ in. knob. A number of shaft-turn-ratios were combined with the levers and knobs. Thirteen subjects completed the series. In over-all performance, the knob with a ratio of 4.08 inches of pointer movement to one complete turn of the shaft proved slightly superior to any of the lever-combinations, with a tolerance of .016 in.—the coarsest that could be obtained with the unmodified apparatus.

Experiment 2. In order to test a much coarser tolerance, a new scale was constructed with .125 in. inserts. With a .025 in. pointer, this permitted a tolerance of .100 in. With a .109 in. pointer, the tolerance was .016 in. as in Experiment 1. Seven subjects, none the same as in Experiment 1, completed the series. With these subjects, there was a slight superiority of levers over knobs at both .016 and .100 in. tolerances.

More striking was the fact that the general level of total times was about two-thirds that of Experiment 1. Are these shorter times simply a matter of differences in the two groups of subjects? Or does the wider insert make for more rapid settings, even when the tolerance is the same?

Experiment 3. The third experiment was designed specifically to answer these questions. A new scale was constructed with alternate wide and narrow inserts (.125 and .016 in.). Alternate runs were made with pointers yielding .016 in. tolerance on both widths of inserts. Six of the seven subjects from Experiment 2 completed the series.

The total times are consistently shorter with the wide inserts (mean difference .015 sec.), although the major difference between Experiments 1 and 2 must have been a difference in the two groups of subjects. It is interesting to note that a wide pointer with a wide insert is better than a narrow pointer with a narrow insert, even though the tolerance is the same.

Experiment 4. To broaden the comparison

* This research was executed under Contract No. W33-038-ac-22561 between the Institute of Research, Lehigh University, and the USAF Air Research and Development Command, Wright-Patterson Air Force Base, Dayton, Ohio.

[1] Detailed tabulated data can be found in USAF Technical Report No. 6563.

between knob and lever combinations, the knob was combined with ratios of 4.08, 6.28, 9.70, and 16.3; while the twelve-inch lever was used with ratios of 16.3 and 33.6. The scale with wide inserts was employed with tolerances of .016 and .100 in. Twenty-four subjects completed the series, making twenty settings apiece at five distances (3, 15, 27, 40, and 72 sixteenths of an inch) under each set of conditions. Thus the over-all mean for each set of conditions is based on 2400 readings.

At a tolerance of .016 in. the knob combined with ratio 6.28 is slightly better than either of the lever combinations. At a tolerance of .100 in. there is little to choose.

Experiment 5. Up to this point no lever longer than twelve inches had been tested. In order to use longer levers, it was necessary to construct a folding arm carrying a double sprocket-and-chain drive; so that the position of the lower end of the lever could be lowered as the lever-length was increased. Lever lengths of 4, 6, 9, 12, 18, 24, and 30 inches were used with ratios of 16.3 and 33.6. Fifteen subjects completed the series. The over-all mean under each condition is based on 1500 readings.

It became evident that the total times were related to the ratio between the movement of the lever-tip (L) and the movement of the pointer (P). The ratio L/P can be computed by the formula:

$$L/P = \frac{\text{lever length in inches} \times 2\pi}{\text{inches of pointer movement for one complete turn of the shaft}}.$$

The main curve of Figure 1 shows the relation between mean total time and L/P. The curve drops rapidly to an L/P of approximately three and then flattens out, possibly rising slightly with higher values. The *optimal* L/P appears to be approximately three or four. That is, the lever tip should travel three or four times as fast as the pointer. By what combination of lever-length and turn-ratio the optimal L/P is achieved seems to make little difference.

For comparison, the mean total time for the knob with ratio 6.28 is shown in the lower left-hand corner of Figure 1. The best performances with knob and with lever come out about even.

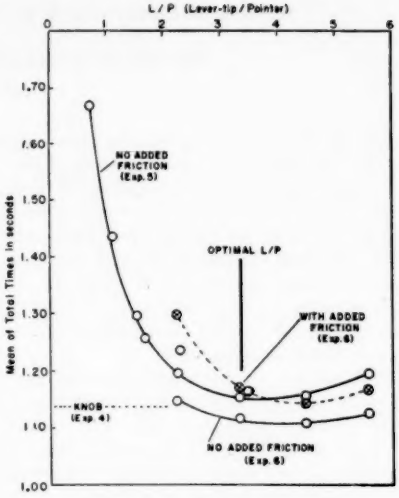

FIG. 1. Total time as a function of L/P.

Experiment 6. By means of a disc-brake, friction was added such as to require an additional 400 grams pull at the tip of the twelve-inch lever. Kinesthetically, this amount of added friction is distinctly noticeable and is described by the subjects as "moderate." Lever lengths of 12, 18, 24, and 30 inches were combined with a ratio of 33.6, giving L/P's ranging from 2.25 to 5.61. The tolerance was .100 in.

The shorter curves in Figure 1 show the relation between mean total times and L/P, with and without added friction. When friction is added, the curve descends more rapidly but the optimal region remains unchanged.

Discussion

For making settings on a linear scale nine inches in length, the optimal L/P appears to be three or four. It seems likely that the same optimum would apply to longer scales, although our apparatus did not permit making such tests.

However, with scales longer than approximately twelve inches, an L/P of three cannot be achieved, because the maximum L is limited by the maximum reach of the subject. As shown graphically in Figure 2, the maximum

is approximately the same (38 in.) regardless of lever length. This means that the maximum L/P cannot exceed 38 divided by the scale length in inches.

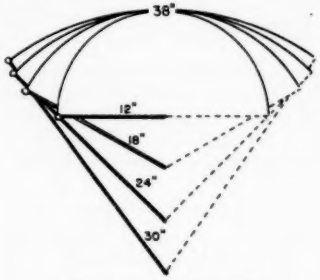

FIG. 2. Maximum reach.

The appropriate shaft-turn-ratio (inches of pointer movement for one complete turn of the shaft) can be computed from the formula:

Shaft-turn-ratio = lever length in inches $\times \frac{2\pi}{L/P}$. For example, with a twelve-inch scale the maximum L/P is $38/12 = 3.17$ which is in the optimal region. If a 30-inch lever is used, the appropriate shaft-turn-ratio is $30 \times 2/3.17 = 60$. If a twelve-inch lever is used, the appropriate shaft-turn-ratio is $12 \times 2/3.17 = 24$.

Summary

Right-left-moving levers of various lengths were used to match a pointer to lighted inserts on a horizontal nine-inch linear scale with tolerances of .016 and .100 in. Rotary control knobs were used for comparison.

With lever lengths between six and thirty inches, the important factor seems to be the ratio of the movement of the lever-tip (L) to the movement of the pointer (P). The optimal L/P appears to be approximately three or four. That is, the lever-tip should move three or four times as fast as the pointer. Adding friction accentuates the optimum without changing it. Slightly faster settings are made with a wide insert and a wide pointer than with a narrow insert and a narrow pointer, even though the tolerance (.016 in.) is the same.

The maximum L is limited by the maximum reach of the subject (about 38 in.). The maximum L/P therefore cannot exceed 38 divided by the scale length in inches. With a scale longer than about twelve inches an L/P of three cannot be achieved.

A rotary knob with a ratio of about six inches of pointer movement for one complete turn of the knob appears to be as good as any lever combination in making settings to coarse tolerances on a nine-inch scale. It seems probable that this would also hold true for longer scales.

Received September 4, 1951.

References

1. Jenkins, W. L., and Connor, M. B. Some design factors in making settings on a linear scale. *J. appl. Psychol.*, 1949, **33**, 395–409.
2. Jenkins, W. L., Maas, L. O., and Rigler, D. Influence of friction in making settings on a linear scale. *J. appl. Psychol.*, 1950, **34**, 435–439.
3. Jenkins, W. L., Maas, L. O., and Olson, M. W. Influence of inertia in making settings on a linear scale. *J. appl. Psychol.*, 1951, **35**, 208–213.

Dimensional Analysis of Motion: III. Complexity of Movement Pattern [1]

Gerald Rubin, Patricia Von Trebra, and Karl U. Smith

University of Wisconsin

The effect of varying the complexity of motion upon the efficiency of skilled movements is a fundamental problem in modern industrial work. But systematic study of the effects of different complexities of movement upon the duration of a motion pattern has never been conducted. The present investigation,[2] which represents part of a program of research on the analysis of work motions, deals with the problem just mentioned, that of the interrelation between complexity of motion pattern and economy of movement.

For experimental purposes, it is essential to define the term, "complexity of movement." Obviously, it is possible to vary the relative difficulty and complexity of a motion pattern by changing the environmental stimuli that effect reactions, the physiological conditions related to the task, or the past learning and other dynamic factors which define the motion. But in terms of the characteristics of the actual movement itself, the relative complexity of motion can be varied in only a limited number of ways. Specifically, the motion complexity may be varied in terms of the following dimensions: (1) the total number of unit movements; (2) the pattern of manipulation; (3) the plane of movement; (4) the directions of movement within a given plane; and (5) the handedness or laterality relations within a repetitive pattern of uniform manipulations of the movement pattern.

In this experiment, the effects of varying the number of directional changes in reactions in a single plane of motion upon the efficiency of a movement pattern have been investigated. New techniques of analysis of human motion, which have been developed for such dimensional study of movements (**2**), permitted variation in the relative number of directional changes in the movement to be made, without altering materially the distance of travel of the motion, the characteristics of manipulation, and other factors which affect the measured properties of the reactions. In addition, the effects of practice on the different complexity-levels of motion used were investigated.

Experimental Procedure

The motion studied in this experiment is a simple work pattern consisting of switch-turning movements used in the operation of a large control panel. The special techniques used in the study permitted separate measures of the basic components of travel time and manipulation time in addition to the over-all performance time. The procedure of the study is to measure simultaneously the travel and manipulation times of the movement pattern under four conditions of complexity.

The apparatus used to conduct this study was the Universal Motion Analyzer (Figure 1). The planned work situation of this particular arrangement of the apparatus consists of a large control panel of separate rows and columns of turn switches. The switches are operated by the subject in predefined patterns. In this apparatus, the subject constitutes a part of the electrical circuit in such a way that his touching and release of successive switches operate two different time clocks. As soon as the subject touches a switch a manipulation time clock is started. When this switch is released, the manipulation time clock is stopped and a travel time clock is started. Upon touching a second switch, the travel time clock is stopped and the manipulation time clock is again started, and so on.

Figure 2 describes diagrammatically the four levels of complexity used in the experiment. In this diagram, the small circles refer to the particular switches turned, and the dotted lines indicate the movements traversed. It should be noted that all four patterns of movement are carried out in the frontal plane. In terms of the preceding definitions, pattern A is the least complex, with zero directional changes, and D is the most complex pattern, with three directional changes. Colored guide strings were used to outline the four different patterns.

Table 1 summarizes the experimental design.[3] The subjects used were 48 right-handed

[1] The cooperation of Omer Jones in the initial conduct of this experiment is appreciated.

[2] Supported in part by funds received from the Graduate Research Committee, University of Wisconsin.

[3] A parallel study of repetitive pin-pulling motions was carried out along with this experiment. The data of this study are not to be published because it was

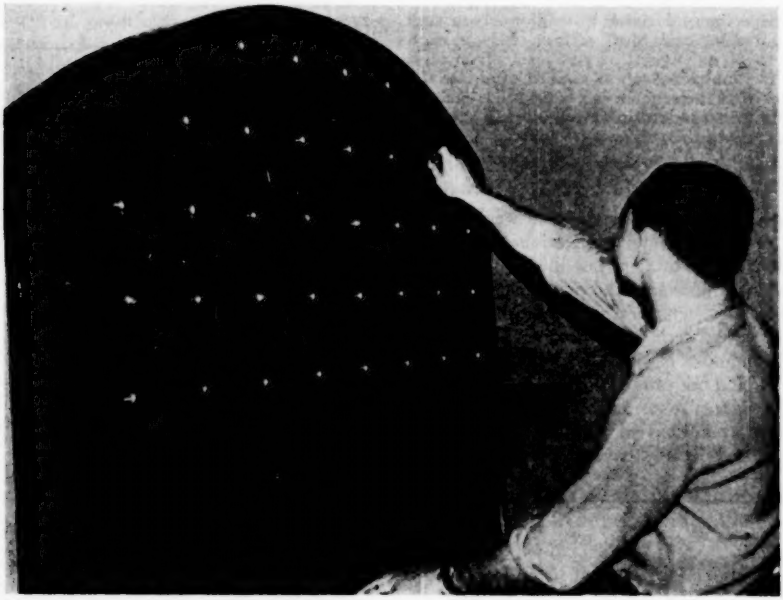

FIG. 1. Photograph of a subject operating a control panel attached to the motion analyzer. The electrical connections between the subject and the apparatus are not shown. Time clocks and electronic units are contained in a separate housing.

college women, who received 4 trials per day for 4 days in each of the 4 patterns of movement. All performances were carried out with the dominant right hand. Order of presentation of the four different complexities of movement was randomized completely by subjects.

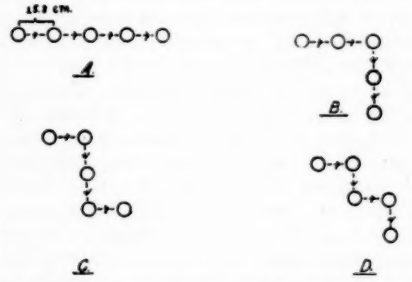

FIG. 2. The patterns of motion used in the experiment.

found that such minor factors as length of fingernails and broken fingernails produced very marked variations in the data. These observations, however, may be of note in connection with the problem of reliability of pegboard psychomotor tests.

The critical data of this experiment are the separate measures of manipulation time and travel time recorded in hundredths of a second for each pattern of movement made by each subject during each trial. In order to treat

Table 1

The Plan of Experimental Observations

Subjects	Trials (16) and Patterns of Movement (4)*	Days** 2	3	4
1	AAAA BBBB CCCC DDDD	–	–	–
2	CCCC AAAA DDDD BBBB	–	–	–
3	DDDD BBBB AAAA CCCC	–	–	–
4	BBBB AAAA CCCC DDDD	–	–	–
.				
.				
.				
48	DDDD CCCC BBBB AAAA	–	–	–

* In the sixteen trials on each day, the order of presentation of the four movement patterns, A, B, C, and D, is randomized by subjects.

** The order of presentation of movement patterns for a given subject is maintained during the four days of practice.

the experimental data, a median score was computed for each block of four trials on each pattern of movement as run on a given day. This median measure was used to eliminate the effects of a few extreme scores caused by blocking and other indeterminate factors in the performances of the subjects.

Results

Separate analyses of variance were carried out for travel and manipulation time data.[4] These analyses brought out the fact that only individual differences and practice effects were statistically significant. Complexity of the movement pattern, which was varied in terms of the number of directional changes in the pattern, did not produce significant differences in the data. Also, none of the interactions between conditions of complexity and other variables, especially days of practice, showed significant variations. The effects of order of presentation of the conditions did not produce significant variation in the experiment or noteworthy influences on other conditions.

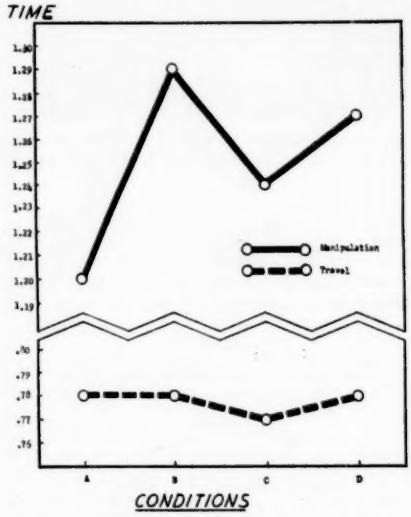

FIG. 3. Manipulation and travel time as a function of conditions of complexity of motion. Time is expressed in seconds. Manipulation time shows some variation with conditions, but neither of the two aspects of movement varies significantly.

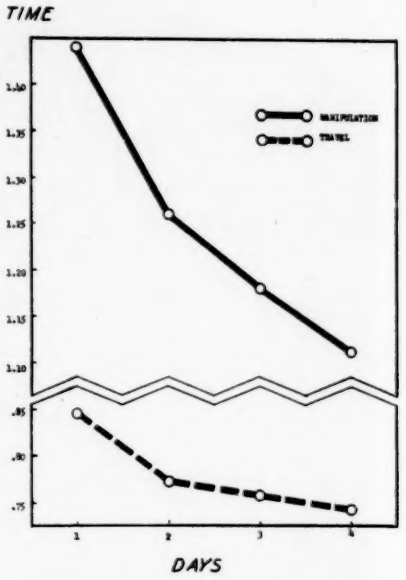

FIG. 4. Manipulation and travel time as a function of days of practice. Time is expressed in seconds. Manipulation time varies significantly with practice.

The concrete results of the experiment may be summarized with reference to some of the actual differences observed. The main object of this study was to determine whether increasing the complexity of a movement pattern, in terms of its directional complexity, will bring about an increase in the time of either the manipulation or travel components of the movement. Figure 3 illustrates the changes in manipulation time and in travel time which were found when the movement complexity was varied in the four conditions. Travel time shows almost no change throughout the four conditions. Manipulation time of the movement pattern varies, but, as already noted, this variation is not large enough to be of any significance. The implication of these results for principles of motion economy will be mentioned later in the discussion.

[4] The summaries of analyses of variance have been deposited with the American Documentation Institute. Order Document 3429 from American Documentation Institute, 1719 N Street, N.W., Washington 6, D. C., remitting $1.00 for microfilm (images 1 inch high on standard 35 mm. motion picture film) or $1.00 for photocopies (6 × 8 inches) readable without optical aid.

A second major objective of this study was to determine the effects of practice on motions of different complexity. Figure 4 illustrates the changes that occur in the manipulation and travel components of motion as a function of practice over a period of four days. As noted above, the effects of practice produce significant variation in the experiment, but this statement applies only to the manipulation component of the motion pattern. The effects of practice on the travel component of the motion are barely significant. As a result of practice, manipulation time in the switch-turning motion is reduced by approximately 30%. Travel time is reduced by approximately 10%, and this change occurs almost entirely during the first day of practice. In other words, this experiment indicates that the effects of learning on the type of work motions studied here have quite distinctive effects on different components of the motion pattern.

In addition to the data obtained concerning the effects of different complexity of movement and of practice, this experiment has provided information concerning the relation between components of movement under the conditions studied. A correlation value has been computed between the travel and manipulation aspects of movement, and has been found to be +0.29. This correlation is significant at the 5% point.

Summary

It has been claimed for many years in the field of motion study that variations in the complexity of a motion pattern bring about loss of efficiency in movement. For example, one of the commonly stated "principles" of motion economy is that the less complex the movement, the more efficient it will be. But no experiments of a systematic nature can be found which, in accordance with meaningful definitions and observations, demonstrate what conditions of complexity produce variations in efficiency of motion. This study points up the necessity of both accurate definition and the obtaining of concrete results in order to understand the conditions of efficient movement. As far as the work motions studied here are concerned, variations in the complexity of a movement pattern in terms of the number of directional dimensions involved do not produce noteworthy changes in the efficiency of the manipulation and travel components of the movement. It remains to be determined by future research what other factors of complexity of motion do influence economy of movement if changes in such complexity do, in fact, define the level of economy of reactions.

A previous study (**3**) has brought out the fact that, in the type of controlled motion patterns studied here, practice does not affect the manipulation and travel components of a motion pattern in the same way, and that the effects of practice do not interact significantly with other conditions of motion. This study therefore substantiates previous work in showing that it is mainly the manipulation component of a motion pattern that is influenced by practice, and not the travel component.

It is believed that the result just mentioned may be a clue to the answer of many of the problems of specificity in skill learning, and of the lack of generality of principles of learning in the practical control and understanding of psychomotor activity. The fact is often recognized that some tasks, for example, tracking behavior, rate of tapping, and throwing activities, typically display very limited practice effects. In these activities, travel or free-thrown movements predominate. Furthermore, the present data point up a fundamental weakness of the traditional generalized approach to the study of psychomotor learning, as well as concepts which have been derived from such an approach. The rules defining the course of learning are quite different for the different primary components of activity in the psychomotor task.

In a previous study (**1**) of a heterogeneous group of right- and left-handed subjects, it was observed that the travel and manipulation characteristics of movement do not correlate significantly with one another. A more homogeneous group of subjects of about the same size, as used in this experiment, has given a correlation of +0.29 between these two aspects of the motion pattern. This low correlation value of limited statistical significance (5 per cent level) suggests that a real problem of

reliable measurement exists in the analysis of psychomotor activity and that this problem may be avoided only by separate measurement of the unrelated components of the work task. The principles and techniques of motion study described here provide the basis for such new approaches to the study of reliability of performance in skilled movements.

Received August 4, 1951.

References

1. Davis, R., Wehrkamp, R., and Smith, K. U. Dimensional analysis of motion: I. Effects of laterality and movement direction. *J. appl. Psychol.*, 1951, **35**, 363–366.
2. Smith, K. U., and Wehrkamp, R. A. Universal motional analyzer applied to psychomotor performance. *Science*, 1951, **113**, 242–244.
3. Wehrkamp, R. A., and Smith, K. U. Dimensional analysis of motion: II. Travel-distance effects. *J. appl. Psychol.*, 1952, **36**, 201–206.

Book Reviews

Cantor, N. *Learning through discussion.* Buffalo 2, New York: Human Relations for Industry, 1951. Pp. 108.

Professional industrial relations personnel who depend upon business and industry for their livelihood face a problem of great magnitude. This problem consists in the main of translating the results of scientific studies in the area of psychology and related fields into language and techniques usable by the layman.

People in the "in-group" are always suspicious of those in the "out-group." The people in the business and industrial world are generally suspicious of what they consider to be "long hairs" and "impractical dreamers." The problem becomes more acute when the representatives of the "out-group" use language that only the most learned can understand and fewer can appreciate.

In addition, most "personnel" men are really very busy. They don't have the time to scan the journals, tie together the various microscopic bits of research and then devise a technique which will have immediate value in the business and industrial situation.

Nathaniel Cantor, the sociologist, has rendered a major service to the professional personnel worker in the publishing of his *Dynamics of Learning* and, more recently, his *Learning Through Discussion*, which is the subject of this review.

Of all the areas in industry and business where the most work needs to be done and where the potential results are the most promising is the area of training. Cantor has provided a practical tool for work in the training area. This tool offers some hope of being understood by his non-technical colleagues and by industrial supervisors and executives.

He states in the preface, "This book deals with the professional discussion leader. It is not another book on 'Conference Leadership' or on 'How to Lead a Discussion' wherein the reader finds listed the things to say or to do. No list of types of questions to ask will be found. There is no 'code' of the conference leader wherein he is supposed to promise himself a series of 'I wills' or 'I will nots.' Plans for preparing a discussion will not be found. There are no check lists.

"The usual manual for leaders contains many such guides which are probably of some value to the beginning discussion leader who feels insecure and uncertain about his function. Listing what he should do and what he should avoid is a poor substitute for lack of knowledge and skill in the art of discussion leadership. Indeed, such devices block the development of personal skill. A concert artist need not be told how to play scales. . . .

"The present approach deals with the fundamental problem, 'What takes place psychologically when a group of people, directed by a leader, meets to discuss a problem?' Another way of stating this is to ask, 'How do members of a group learn or change, by the exchange of opinions through argument, or by discussion on an intellectual, verbal basis? Is learning a matter of reasoning or does it involve much more?'

"It is my firm conviction that a discussion leader cannot perform in a professional way unless he understands what is involved in learning. To acquire such understanding and the skill to employ it are no simple matters. Ideally, there should be instituted a professional program for the development of discussion leaders. Until such time, we will have to rely upon the meager literature on discussion leadership and the unsatisfactory and hurried 'in-service' programs."

This approach is where the major contribution of the book lies. It presents techniques for the application of sound learning without bogging down in technical language and formulae. It provides the professional training director with a readily usable tool for the training of discussion leaders. It is this non-professional teacher who makes possible a sound educational program which can be continued on an economical basis.

Cantor's use of Kelley's basic educational fallacies to develop the chapters on Learning is excellent. It strikes home and is appealing to anyone who has "escaped" from our public school system.

The chapters which follow flow naturally

into the discussion of the Dynamics of Discussion and the Skills in Discussion. He points out, "Most of the members of a discussion group are not likely to be disciplined social scientists. They are not accustomed to consider issues in terms of the logic of ideas. They bring to the discussion their stereotyped 'rights' and 'wrongs,' 'goods' and 'bads,' their 'either-or' views of social behavior. They are not in the habit of examining the hidden assumptions or biases in their points of view. They interpret social events and individual behavior in terms of their more or less fixed attitudes. They tend to confuse the observation of premises with their moral judgments about them.

"One of the responsibilities of the discussion leader is to help them learn to look at events from many different and conflicting points of view. As new facts and new interpretations arise, previous attitudes may be modified."

It is this type of presentation of psychological truths in simple language that makes the book so useful. If more of the body of research which has been published in group dynamics, leadership, counseling and psychotherapy can be translated into easily used technique and presented in simple language in the way Cantor has done, the business and industrial world would stand up and cheer.

Howard P. Mold

Director of Training,
Minneapolis-Honeywell Regulator Company

Thorne, F. C. *Principles of personality counseling.* Brandon, Vermont: Journal of Clinical Psychology, 1950. Pp. 491.

For the most part, this book represents a collection of the series of articles on counseling and psychotherapy published by Dr. Thorne over a number of years, mostly in the *Journal of Clinical Psychology.* In most cases the articles have been modified or combined to conform to the organization of the book. Here and there transitional and integrating discussions have been inserted. Thus, this work does not represent a new contribution to the literature, but does provide a more effective vantage point for an over-all view of Dr. Thorne's contributions.

For the counselor it has certain definite though limited values. In so far as there is a trend toward interpreting counseling as a clinical process designed to serve broad preventive and positive mental health functions, the broad range of treatment problems discussed with the emphasis on the borderline neurotic and relatively normal individual will meet a growing need. On the other hand, many of the discussions and cases are oriented toward problems where psychological and physiological factors are closely intertwined. These aspects of the book will have more immediate application in psychiatric practice or in psychological practice in medical agencies than in the usual practice in psychological counseling. Particularly illustrative of this limitation is Chapter 3 in which discussions of testing and diagnosis are limited to evaluation of psychopathology. All counselors, including those who confine themselves to helping people make decisions, will obtain food for thought from the many emphases on cognitive processes in helping people with personal problems, e.g., imparting information, perceptual training, and dealing with memory. The last two sections are devoted to maximizing intellectual resources and to methods of intellectual reorientation, respectively. Too often, an interest in "depth" factors in personality leads to unwise neglect of cognitive intellectual elements in personality adjustment.

The author is long on discussions of practical problems, almost fanatical in expressing his devotion to science and eclecticism, but short in the theoretical coherence of his various concepts. This reviewer had the impression that it would be easy for readers to acquire a series of undigested somewhat contradictory ideas about counseling and other therapeutic phenomena as a consequence of the book's rampant eclecticism. On the other hand, discussions such as those of practical problems of professional responsibility (chapter 5) and of creating an effective counseling relationship (chapter 8) will prove useful to many beginning counselors.

In summary, many parts of this book can provide useful supplementary reading for didactic courses in counseling.

Edward S. Bordin

University of Michigan

Hathaway, S. R., and Meehl, P. E. *An atlas for the clinical use of the MMPI.* Minneapolis: Univ. of Minnesota Press, 1951. Pp. xliv + 799. $9.75.

Despite the title, the *Atlas* does not deal with clinical uses of the MMPI. It consists of 798 brief case histories abstracted by the senior author from the clinical records of in-patients at the Psychiatric Unit of the Univ. of Minnesota Hospitals. These are supplemented by 170 cases, obtained from eleven other sources, including prisoners, college students, Veterans Administration hospital patients, guidance clients, and patients in an English hospital. The histories are factual rather than interpretive in nature.

Each of the 968 cases is headed by one or more MMPI profiles and related diagnostic and descriptive data. The profiles have been reduced to a code which summarizes the form or shape and gives some information about the intensity or elevation. The cases are arranged according to this code and are extensively indexed and cross-indexed. This enables the user to look up cases on the basis of the MMPI pattern and provides material on all the usual profile configurations and many atypical ones.

However, the reader will have to have had considerable experience with the MMPI before he can profitably use the *Atlas*. The discussion prerequisite to an informed and sophisticated use of the instrument is not provided in the book. For those unfamilar with the literature a 259-item bibliography (through November, 1950) is given and sixteen articles are starred as recommended reading.

There are eight figures showing in graphic form the relative frequencies of two-digit codes for both male and female psychiatric patients, normal adults, college students, and ninth-graders. Eight tables show for the same groups the frequencies of all combinations of highest and lowest scales in the code. Of particular interest are other tables giving two-digit codes for fifteen common diagnostic groups contrasting their relative frequencies with those obtained in a normal population and in a general clinical sample.

There are a great many relationships inherent in the tabular and textual material which the authors have not made explicit. It may be expected that other workers will carry out researches on the contents of this volume much as governmental and census statistical reports are often utilized. This would require evidence for the accuracy and representativeness of the samples and case material presented; such evidence is, unfortunately, at present not available for evaluating the book from this standpoint.

The coding employed in the *Atlas* has two major disadvantages: first, scales with T-scores from 46 to 55 are not coded; second, "low" scales with scores below 46 are coded from lowest to highest—just the opposite of the "high" scales. This reversal led to awkwardness if there is more than one low scale. The first deficiency causes variability in code length if any scales lie in the uncoded range. Also it may be impossible to tell which is the lowest (or in some cases the highest) scale if more than one scale is uncoded. This difficulty is exemplified by Table I, page xxvii where 50.4% of 710 male psychiatric patients have their low points uncoded; there is no way of determining whether these low points are distributed in the same proportions as those with coded low points. The utility of all the tables and figures is reduced because of this shortcoming.

The *Atlas* should encourage clinical workers to utilize the profile patterning and configural approach to the MMPI and should discourage the unprofitable adherence to the diagnostic terminology of the individual scales. It will then be possible to determine empirically and without psychiatric bias the personality correlates—both normal and abnormal—of the various profile patterns.

George S. Welsh

Veterans Administration Hospital, Oakland, California

Read, Herbert. *Education through art.* N.Y.: Pantheon Books, 1949. Pp. xxiii + 320. $5.50.

This book will interest psychologists principally because it sets forth an eclectic theory of personality development through creative expression. Read develops an idea as old as Plato—that art should be the basis of any

sound education—and offers it to the modern world as a solution for society's ills.

Read's material on sensation, perception and imagery represents classical psychology. It bears the distinct impress of British associationism, shows the color of Gestalt theory, but has little mark of American behaviorism or neo-behavioristic viewpoints. His personality theory is chiefly based on types which bear the hallmarks of Jung and Jaensch; he tends strongly to speak in terms of dichotomies.

Read's position on art in relation to thinking, discipline, and moral training illustrates several diffuse but pronounced trends in modern education. It is antirationalistic, emphasizing an intuitive approach to truth. Certain aspects of experience exist in the nature of things (e.g., balance, symmetry, proportion, rhythm), and are apprehended because they "feel right" to the observer. The tendency to empathy and abstraction in art is deeply rooted in human nature. The artist, not the logician, is close to the personality and mind of the child. Education consists in nurturing the relatively pure tendencies of the child mind and avoiding the harm which comes through ordinary educational procedures.

Read's theory also holds that productive thinking in the scientific sense is closely related to artistic production and may be governed by the same principles. The elements of intuition, discovery, "insight" are given an important place. Today, when many persons are becoming concerned with the nature of scientific creativity, this antirationalistic definition of scientific thinking is more than a curiosity. One notes many kindred viewpoints in modern thought. Whether the psychological reader agrees with Read's position or not, he might well try to discover why such a position recurs and why educators appear receptive to what in earlier years was considered so definitely unscientific.

Read derives a series of categories for the evaluation of child art and advances a theory of art production in relation to human personality which will interest those concerned with projective techniques. He also offers a blueprint for adapting his educational and personality theories to the education of children and adolescents.

Dale B. Harris

Institute of Child Welfare, University of Minnesota

Link, Henry C. *The way to security.* New York: Doubleday and Company. 1951. Pp. 224. $2.50.

In an article written in 1939, I quoted H. A. Murray to the effect that the kind of answers that satisfy people's curiosity and meet their psychological needs was being provided not by psychologists but by medical men. To make the observation adequate today, one must add to medical men, preachers, priests, rabbis, and chautauqua lecturers. Henry C. Link, however, is one psychologist who has consistently written books that satisfy people's needs. Book reviews in newspapers and magazines and sales records testify to the popularity of *The Return to Religion, The Rediscovery of Man,* and the *Rediscovery of Morals.* They have not been equally satisfying to his fellow psychologists for at least two reasons: (1) they are anecdotal in style; and (2) they describe and explain in terms of the Spirit and things spiritual, entities which are not encountered in a life time of work in the laboratory.

Link's most recent book, *The Way to Security,* is in similar vein and promises to satisfy the one group and dissatisfy the other as before. Its title reflects the keen sense of the marketing expert, but its content derives from a long experience as psychologist, philosopher, father of a family, personal counselor and sympathetic observer of people. As in his earlier books, he is in revolt against present day materialism.

Link notes that with all the current provisions for security by society and by the individual, the sense of insecurity was never so great. The reason is that security has come to mean social security and material security whereas the basic and satisfying security is personal and spiritual. The latter cannot be derived from the former. Why has the search for security shifted from the personal and spiritual to the material and social? Why has security come to be something to be *received* rather than something to be *achieved?* Link

finds the answers to these questions in twenty-seven pseudo-scientific fallacies, whose discussion and illustration constitute the content of his book. He finds these fallacies in the home, in the church, in the schools, and in the government. These fallacies lead to the building up of an insecure personality devoid of self-reliance.

Link goes to the dictionary for a definition of security and finds that "to be secure" means "to be fastened to." Taking off from that meaning he shows that young people today have nothing to fasten on to as earlier generations did. The old fixed principles of right and wrong, the ten commandments and the Judaeo-Christian ethics have been replaced by the concept of "let your conscience be your guide." To bewildered parents who ask him to recommend a book on how to bring up children he usually replies: "If you will do with your children pretty much what your parents and grandparents did with theirs, it will be better than any modern book I can recommend. It may not be perfect but at least it will not be bad" (page 50).

Concerning religious faith he says: "The basis for authority and security is a faith in certain moral principles and standards. That faith has been severely shaken, bringing about a calamitous decline in discipline. Now an ever-growing mass of evidence proves that the faith of our ancestors was not entirely misplaced. Neglected principles that were centuries in the making are being rediscovered and reinterpreted in terms of modern life" (page 56).

Link's most severe criticism against formal education is that though it educates the intellect it does not develop personality. If students wish "to develop mature character they must rely on their families and the church. Inadequate though these may be, they are still far superior to the forces of public education. Indeed, since the training of character must be largely based on religious principles, the public schools cannot do this training properly even if they would. Religious education is, for certain reasons, no longer compatible with secular education" (page 135). Physical education seems to Link to offer the best medium for spiritual education within the modern educational system (page 138).

The Welfare State comes in for severe indictment because it is directed toward material and social security with the impoverishment of personal and spiritual security. ". . . a government cannot assume responsibility for people's welfare without profoundly affecting their moral fibre. To the extent that government takes care of him, to that extent the adult citizen is deprived of the moral responsibility for himself. And since a government without religion tends to be neutral in its moral standards, it cannot make such standards a condition of its welfare payments. The net result is the progressive breakdown in the moral standards of all who participate in the welfare state" (page 224).

Space does not permit comments on the many specific topics that Link discusses and illustrates from his extensive counseling experience, such as: the need for love, authority and discipline; sex, love, marriage and children; the psychology of fear and worry; sleep and relaxation; the mature character; war and security; the threat of communism from within; and the dollar as a gauge of spiritual security.

The last few sentences of the book give at once the causes underlying a sense of insecurity and the remedy for it. They are: "By one false theory after another we have emasculated the concept of God to its present innocuous stature. Still further, we have created God in our own image and can now contemplate him without a sense of fear or sin. Such complacency is not conducive either to spiritual or material security.

"Nevertheless the good news that God came to save man because man cannot save himself brings increasing conviction. And with this conviction also comes the certainty that man can be saved if he will make the effort to meet God halfway" (page 224).

A. T. Poffenberger

Montrose, N. Y.

Institute for Human Adjustment. *Training of psychological counselors. Report of a conference held at Ann Arbor, Michigan, July 27 and 28, 1949, and January 6 and 7, 1950.* Ann Arbor: University of Michigan Press. Pp. 29.

This significant document is the report of a conference held at the University of Michigan

to discuss "the training of counselors whose background is to be primarily psychological." The conference was planned by the Executive Committee of the APA Division of Counseling and Guidance and the University of Michigan Bureau of Psychological Service, with the help of the National Institute for Mental Health, USPHS. Twelve persons met in the two conference periods, representing the National Institute for Mental Health, the U. S. office of Education Division of Occupational Information and Guidance, Clinical Psychology, an industrial counseling agency, and the APA Division of Counseling and Guidance. Care was taken in sampling from the last group to represent proponents of "informational and general student personnel procedures" as well as those favoring "objectives of emotional and personality development" (p. 4).

The conference procedure and report preparation (see pp. 4–5) seem to have permitted a useful synthesis of participant contributions. Perhaps the most valuable outcome is the description of psychological counseling and of levels and content of training. Psychological counseling is viewed as a generic method of helping an individual who has problems arising from his interpersonal relations to (1) increase the accuracy of his "self-percept," (2) increase the accuracy of his "environmental perceptions," (3) integrate "self-percepts" and "environmental perceptions," (4) obtain relevant information, and (5) improve in his capacity to plan and execute behavior effectively outside of the counseling situation. As might be expected in a preliminary statement, much of this material needs to be revised. By what criteria does one test whether one is "increasing the accuracy of the individual's self-percept" (p. 7) or "aiding the individual to make a better—in the sense of happier—adjustment to his life situation" (p. 10)?

Recommended levels of training include: (1) the "part-time counselor" with one year of graduate work, two-thirds of which would be in psychological training, (2) the "psychological counselor" with two years of graduate work, primarily psychological, approximately one-third of which would be in counseling theory and practice, and (3) the "counselor-psychologist" (why not "counseling psychologist"?) who would be trained at the doctoral level. The conference report is particularly helpful in describing training and standards for Levels (1) and (2). The doctoral program is only briefly discussed.

Of course, some important issues have not been resolved. For example, one infers that there was sharp dissension over the requirement of teaching experience prior to the practice of psychological counseling in an educational setting (p. 16). There does seem to have been agreement, however, that for the counselor in training there should be "some members of the teaching staff with adequate background in both theoretical training and practical experience" (p. 24), and "that no one should attempt to counsel independently until he has had a period of closely supervised experience" (p. 18).

Exception may be taken to "the prevailing opinion . . . that formal learning theory was not at the stage where it had any direct contribution to the training of this type of applied worker" (p. 17). This flat pronouncement is unfortunate in the face of the many relevant and seemingly helpful attempts to integrate learning hypotheses with counseling procedure which have been provided in recent years by such writers as Dollard and Miller, Mowrer, Magaret, and Shoben.

In conclusion, the participants—especially E. S. Bordin, conference chairman and report editor—are to be congratulated upon their preliminary job of definition and prescription. The report should help to bring about a needed clarification of counselor activities and training within the Division of Counseling and Guidance, APA, and among related professional groups who may perform "psychological counseling" functions.

Harold B. Pepinsky

The Ohio State University

Boder, David P. *I did not interview the dead.* Urbana: University of Illinois Press, 1949. Pp. xx + 220. $3.50.

Initially, Dr. Boder's project seems fascinating. In 1946 he travelled to several of the European displaced persons camps, electrically recording his interviews with the inmates concerning their recent experiences. As a group, these DP's have gone through a set of

conditions more extreme, more potentially destructive of adjustment patterns than any to which a sizeable group of humans in recent history has been exposed. The plan of obtaining research material from them in this form is potentially profitable. But to be worth while, the venture must satisfy a few basic requirements in aims, sampling and controls. Similarly, this volume as a report of his work must meet other requirements to make any scientific contribution.

The report does not state any definitive purpose other than a general one of getting research material. If this reflects the actual lack of guiding hypotheses in the collection of the data, then it is difficult to see how the material can serve adequately for either a nomothetic or an idiosyncratic analysis. To use the information to characterize some group the analyst needs some description of the sampling model used. He reports only, "In all, one hundred and twenty hours of interviewing were recorded, covering stories of about seventy people, representing nearly all creeds and nationalities in the DP installations in the American Zone. Seven complete narratives and parts of one, the longest interview, are reported in this book" (p. xiii). Of these latter, six are Jewish, one Mennonite and one Greek Orthodox. What generalizations can be formed on these selections?

If the author is offering this material only for its value in generating hypotheses, then he seems to have defeated his purpose in at least two ways. First of all, in the meager amount of material printed in this volume, since the whole seventy cases themselves would have been very thin fare in contemporary sociological research. Secondly, he mentions rejecting some cases because their experiences seemed too unusual, desiring instead the rank and file DP's. Often in an exploratory survey, however, it is the extremes which are desired to set some limits to the range of phenomena in the domain.

Examination of the available cases casts serious doubt on the worth of the unpublished material if he did not control the uniformity of coverage any better on the others. Most of the common principles of obtaining testimony have been violated. In rejecting what he calls "the usual 'pencil and paper' method of interview," he has thrown out most of the measures designed to assure adequate coverage of material in all cases. As a result, when the reader attempts to apply Dr. Boder's "Traumatic Index," which is given in an abridged form only, he finds it impossible because the cases are not sufficiently comparable.

Nothing is known of the cases but what is obtained in the interview. Consequently no checks on the range of possible errors of this method of collection of data are possible. Conceivably, there does remain the possibility of intra-interview analysis, treating the recording as a personal document. The author suggests this possibility when he mentions evidence of traumatic disruption of their language habits, or the presence of inconsistencies within the records. Here too the analyst is frustrated by the clumsy and haphazard interviewing methods used. Time and again, the subjects start relating in their own way some personal experience only to be interrupted by seemingly trivial or irrelevant queries, asides or equivocations. These are so frequent as to distort seriously the whole set of documents.

These and similar errors are sufficiently damaging to detract from the worth of this book for psychology, sociology or anthropology. Perhaps this was not the actual aim of the author's efforts, however. A key to the underlying motivations seems to lie in the introductory paragraph on page xi where he mentions the "significance of preserving for posterity the impressions and emotions *aroused* by the sight of thousands of victims . . ." (italics reviewer's). What is most valued is not the behavior of the subjects interviewed, but the behavior of the audience in its reaction of outrage. As such it is not scientific study but propagandistic efforts at arousing emotion. If this is so, it should be clearly labelled as such, and evaluated as such. Even as propaganda, although it is powerful, it does not come up to John Hersey's *The Wall*.

W. Grant Dahlstrom

State University of Iowa

New Books, Monographs, and Pamphlets

Books, monographs, and pamphlets for listing and possible review should be sent to Donald G. Paterson, Editor, Department of Psychology, University of Minnesota, Minneapolis 14, Minnesota.

The psychology of teaching reading. Irving H. Anderson and Walter F. Dearborn. New York: The Ronald Press Co., 1952. Pp. 380.

The public librarian. Alice I. Bryan. New York: Columbia University Press, 1952. Pp. 474. $6.00.

Applied psychology. Abridged edition. Harold E. Burtt. New York: Prentice-Hall, Inc., 1952. Pp. 480.

Practical applications of democratic administration. Clyde M. Campbell, editor. New York: Harper and Brothers, 1952. Pp. 325. $3.00.

Factor analysis. Raymond B. Cattell. New York: Harper and Brothers, 1952. Pp. 462. $6.00.

Counseling in catholic life and education. Charles A. Curran. New York: Macmillan Co., 1952. Pp. 462.

Proceedings of the international workshop on guidance. Mitchell Dreese *et al.* Germany: Office of Public Affairs, Education and Cultural Relations Division, Education Branch, Office of the United States High Commissioner for Germany, 1951. Pp. 171.

A basic text for guidance workers. Clifford E. Erickson, editor. New York: Prentice-Hall, Inc., 1952. $4.75.

The counseling interview. Clifford E. Erickson. New York: Prentice-Hall, Inc., 1952. $2.50.

Social psychology. Robert E. L. Faris. New York: The Ronald Press Co., 1952. Pp. 420.

Understanding heredity. Richard B. Goldschmidt. New York: John Wiley and Sons, Inc., 1952. Pp. 228. $3.75.

Psychology in industry. J. Stanley Gray. New York: McGraw-Hill Book Co., Inc., 1952. Pp. 401. $5.00.

Measurements of human behavior. Revised edition. Edward B. Greene. New York: The Odyssey Press, Inc., 1952. $4.75.

Growing through play. Ruth E. Hartley. New York: Columbia University Press, 1952. Pp. 62. $.75.

New play experiences for children. Ruth E. Hartley, Lawrence K. Frank, and Robert M. Goldenson. New York: Columbia University Press, 1952. Pp. 66. $.75.

Readings in industrial and business psychology. Harry W. Karn and B. von Haller Gilmer. New York: McGraw-Hill Book Co., Inc., 1952. Pp. 478. $3.50, paper; $4.50, cloth.

Increasing personal efficiency. Donald A. Laird. Fourth edition. New York: Harper and Brothers, 1952. Pp. 294. $3.50.

Incentive management. James F. Lincoln. Cleveland: Lincoln Electric Co., 1951. Pp. 280. $1.00.

The yearbook of psychoanalysis. Sandor Lorand, managing editor. New York: International Universities Press, Inc., 1952. Pp. 271. $7.50.

Marketing research. David Luck and Hugh Wales. New York: Prentice-Hall, Inc., 1952. Pp. 480.

Art and technics. Lewis Mumford. New York: Columbia University Press, 1952. Pp. 162. $2.50.

Lectures and conferences on mathematical statistics and probability. Jerzy Neyman. Second edition. Washington, D. C.: Graduate School, United States Department of Agriculture, 1952. Pp. 274. $5.00.

Play, dreams and imitation in childhood. Jean Piaget. New York: W. W. Norton and Co., Inc., 1952. Pp. 296. $5.00.

Lines in progress, a study of the natural growth of personality. Robert W. White. New York: William Sloane Associates, Inc., 1952. Pp. 320. $2.75.

The hand in psychological diagnosis. Charlotte Wolff. New York: Philosophical Library, 1952. Pp. 218. $7.50.

Personality and problems of adjustment. Kimball Young. Second edition. New York: Appleton-Century-Crofts, Inc., 1952 Pp. 294. $5.00.

Articles on

READABILITY FORMULAS

REPRINTED FROM THE

Journal of Applied Psychology

"A New Readability Yardstick," by Rudolf Flesch.

"Tables for Use with the Flesch Readability Formulas," by James N. Farr and James J. Jenkins.

"Simplification of Flesch Reading Ease Formula," by James N. Farr, James J. Jenkins, and Donald G. Paterson.

•

Prices:

Each reprint 25¢. A set of the three reprints 50¢.

Quantity orders:

25 reprints $3.50	25 sets $7.00
50 reprints $5.00	50 sets $10.00

•

Order from

AMERICAN PSYCHOLOGICAL ASSOCIATION

1515 Massachusetts Ave. N.W.

Washington 5, D. C.

JOURNAL OF APPLIED PSYCHOLOGY

YEAR	VOLUME	AVAILABLE NUMBERS						PRICE PER NUMBER	PRICE PER VOLUME
		MAR	JUN	SEP	DEC				
1917	1	–	–	3	4			$1.75	$3.50
1918	2	1	2	3	4			$1.75	$6.00
1919	3	1	2	3	4			$1.75	$6.00
1920	4	1	2 &	3	4			$1.75	$6.00
1921	5	1	2	3	4			$1.75	$6.00
1922	6	1	–	–	4			$1.75	$3.50
1923	7	–	–	3	4			$1.75	$3.50
1924	8	1	2	3	4			$1.75	$6.00
1925	9	1	2	3	4			$1.75	$6.00
1926	10	–	2	–	–			$1.75	$1.75
		FEB	APR	JUN	AUG	OCT	DEC		
1927	11	1	2	3	4	5	6	$1.25	$6.00
1928	12	1	2	3	4	5	6	$1.25	$6.00
1929	13	1	2	3	4	5	6	$1.25	$6.00
1930	14	1	2	3	4	5	6	$1.25	$6.00
1931	15	1	2	3	4	5	6	$1.25	$6.00
1932	16	1	2	3	4	5	6	$1.25	$6.00
1933	17	1	2	3	4	5	6	$1.25	$6.00
1934	18	1	2	3	4	5	6	$1.25	$6.00
1935	19	1	2	3	4	5	6	$1.25	$6.00
1936	20	1	2	3	4	–	6	$1.25	$6.00
1937	21	1	2	3	4	5	6	$1.25	$6.00
1938	22	1	2	3	4	5	6	$1.25	$6.00
1939	23	1	–	3	4	5	6	$1.25	$6.00
1940	24	–	–	–	4	–	6	$1.25	$2.50
1941	25	–	2	3	4	5	6	$1.25	$6.00
1942	26	1	2	3	4	5	6	$1.25	$6.00
1943	27	1	2	3	4	5	6	$1.25	$6.00
1944	28	1	2	3	4	5	6	$1.25	$6.00
1945	29	1	2	3	4	5	6	$1.25	$6.00
1946	30	1	2	3	4	5	6	$1.25	$6.00
1947	31	–	–	–	4	5	6	$1.25	$3.75
1948	32	1	2	3	4	5	6	$1.25	$6.00
1949	33	1	2	3	4	5	6	$1.25	$6.00
1950	34	1	2	3	4	5	6	$1.25	$6.00
1951	35	1	2	3	4	5	6	$1.25	$6.00
1952	36	By subscription, $6.00						$1.25	
								List price, Volumes 1 through 35	$192.50
								30% Discount	57.75
								Net price, Volumes 1 through 35	$134.75

Information about the Journal of Applied Psychology: 19 numbers are out of print, according to the inventory of December 31, 1951.

Volumes 1 through 10 had four numbers per year; the remaining volumes have six numbers per year.

Information about prices: the journal is uniformly priced at $6.00 per volume. When four numbers per volume were issued, the price per number is $1.75; when six numbers per year were issued, the price per number is $1.25. For foreign postage, $.25 per volume should be added. The American Psychological Association gives the following discounts on orders for any one journal:

10% on orders of $ 50.00 and over
20% on orders of $100.00 and over
30% on orders of $150.00 and over

Current subscriptions and orders for back numbers should be addressed to

AMERICAN PSYCHOLOGICAL ASSOCIATION
1515 Massachusetts Avenue N.W.
Washington 5, D. C.